Interindividual Differences in Verbal Irony Detection and Use:

The Role of Personality and Ability

Thesis (cumulative thesis)

Presented to the Faculty of Arts and Social Sciences

of the

University of Zurich

for the Degree of Doctor of Philosophy

by Richard Bruntsch

Accepted in the spring semester 2017

on the recommendation of the doctoral committee:

Prof. Dr. Willibald Ruch (main supervisor)

Prof. Dr. Karl-Heinz Renner

(Zurich, 2017)

Acknowledgments

Over the past few years, the work I have conducted for this PhD thesis has benefitted immensely from many people's support and inspiration. My gratitude goes to all the people who helped me to strike root in the field of personality and assessment, and eventually made it possible for my work to bear fruit, not least in the form of this thesis.

I am grateful for the professional guidance provided to me by Prof. Dr. Willibald Ruch, who knowledgeably directed me to the "nourishing sources" that helped my work to develop and who taught me how to "tell a good story" in a research report.

I wish to thank Prof. Dr. Karl-Heinz Renner for his supervision and for giving me essential advice and directions for my PhD work.

I am indebted to Dr. Jennifer Hofmann for sharing her academic experience, skills, and knowledge with me and for being as helpful as a research mentor can be.

My work benefitted from the fertile soil of expertise and skillfulness in the department of personality and assessment of the University of Zurich, where it developed on top of the foundation of personality and humor research that the department had laid before my time as a PhD candidate. Furthermore, without the professional assistance of and the enjoyable comradeship with my fellow researchers, my time as a PhD student would have certainly been less fruitful and gratifying.

I would like to acknowledge and thank the *Stiftung Humor und Gesundheit* for financially supporting me in covering the cost for participants' incentives to take part in the studies.

I would like to thank those people who took the time to read and knowledgeably critique my work, such as my supervisors and peer reviewers, as I learned a lot from this.

I wish to express my gratitude to my family and friends for taking interest in my work and, most of all, for enriching my time as a PhD candidate beyond the professional domain during the past few years.

Abstract

Being the first to ask the question whether (a) systematic interindividual differences in irony performance can be found and (b) whether this interindividual variance can be explained by personality and ability variables, the present thesis aims to paint a clearer picture of *who* is able or inclined to detect or use irony. The aim in dealing with these questions is to open up a new field of study for both personality and irony research by conceptualizing irony detection as an aptitude and irony use as an enduring tendency. The results support these expectations by demonstrably (a) linking the aptitude to detect irony to general mental ability and personality traits, and (b) linking the tendency to use irony to personality traits. More broadly, the results support the central claims in that they indicate (a) that there is systematic and measurable interindividual variance in irony detection and use, and (b) that a substantial amount of this interindividual variance can be explained by variables from the realm of ability and personality as two central domains of individual differences. Furthermore, the present thesis introduces a new means of assessing irony detection performance (i.e., including ironic praise as a previously neglected category of stimuli). As a secondary result, the present thesis also supports pre-existing assumptions about the role of humor in irony behaviors. The findings have implications for different fields of irony and humor research.

Zusammenfassung (German abstract)

Diese Arbeit geht der Frage nach (a) ob sich systematische interindividuelle Unterschiede in Ironieverhalten nachweisen lassen und (b) ob sich diese interindividuelle Varianz mittels Persönlichkeits- und Fähigkeitsvariablen aufklären lässt. So zielt die vorliegende Arbeit darauf ab, ein deutlicheres Bild davon zu gewinnen, wer dazu fähig oder geneigt ist, Ironie zu entdecken bzw. zu gebrauchen. Ziel der Arbeit ist es, ein neues Feld sowohl in der Ironie- als auch in der Persönlichkeitsforschung zu erschliessen, indem Ironieentdeckung als Fähigkeit und Ironiegebrauch als überdauernde Eigenschaft aufgefasst wird. Die Ergebnisse stützen die getroffenen Annahmen, insofern als empirisch (a) die Fähigkeit, Ironie zu entdecken mit Intelligenz und Persönlichkeit und (b) die Neigung, Ironie zu gebrauchen mit Persönlichkeit in Verbindung gebracht werden konnte. Die Ergebnisse deuten darauf hin, dass (a) es systematische und messbare interindividuelle Varianz in Ironieentdeckung und –gebrauch gibt und (b) ein wesentlicher Anteil dieser Varianz durch Variablen aus dem Bereich der Fähigkeiten und Persönlichkeitsmerkmale aufgeklärt werden kann. Darüber hinaus stellt die vorliegende Arbeit ein neues Verfahren zur Messung von Irnonieentdeckung vor, das auch ironisches Lob einschliesst. Ferner stützt die vorliegende Arbeit auch bestehende Annahmen über die Rolle von Humor bei Ironieverhalten. Die Ergebnisse haben eine Tragweite für verschiedene Bereiche der Ironie- und der Humorforschung.

Contents

List of Tables

List of Figures

General Introduction

The present thesis has three broad goals: (a) to explore whether irony detection and use can be found as individual differences phenomena, i.e., whether there is systematic interindividual variance in terms of an irony detection aptitude and an enduring tendency to use irony, (b) to derive and test hypotheses as to which individual differences variables play a role in irony detection and use, and, as a secondary aim, (c) to contribute to the literature by ultimately deriving a theory that—in terms of the individual differences variables found or assumed as facilitating or impeding irony behaviors—explains why there are differences between individuals with regard to the extent to which they successfully receive irony.

Irony detection has been extensively investigated in terms of comparing mean performances between experimental groups (e.g., Gibbs, 1986; cf. Kreuz, 2000), and also between healthy control subjects and natural groups (such as clinical groups) known as having impaired cognitive prerequisites for detecting irony (see section "Clinical groups known for *Theory of Mind* deficits"). These findings illustrate that there are known variables associated with limitations in cognitive and emotional inference processes that have an impact on verbal irony comprehension. When looking at the variance within the compared groups, these experimental studies also provide indications for the assumption that there is variance in irony detection that can be measured, even among healthy adults (see PART I). This is important, because without variance in the phenomena investigated, assuming an individual differences perspective would not make any sense. In fact, as a basic assumption to be studied in the present thesis, it is expected that there is a substantial amount of interindividual variance in irony detection performance. Most notably, it is assumed that these interindividual differences are *systematic* and *meaningful* (rather than negligible "noise") as irony performance (a) is expected to *correlate with itself* (in terms of a stable and

enduring aptitude) and (b) is hypothesized to correlate with theoretically relevant traits and abilities.

Likewise, there are only very few studies on irony *use* in general and most of them have investigated cultural or situational factors predicting irony use (see PART III). Only very few studies that investigate irony detection and use from an individual differences perspective can be found. These studies also support the assumption that (a) there is interindividual variance in irony detection and use; furthermore, they exemplify that (b) this variance can be explained by individual differences variables, that is, in terms of personality traits and ability. However, these studies focus only on a few and rather narrow traits, such as when linking irony detection to *postformal thinking* (Blouin & McKelvie, 2012) or *need for consistency* (Groeben, Seemann, & Drinkmann, 1985). As a further example, Langdon and Coltheart (2004) found that the schizotypal personality trait is associated with irony detection, suggesting that individuals who are assumed to have a heightened risk for psychotic illness (i.e., high scorers in the schizotypal personality measure) have an impaired aptitude to detect irony. As to irony *use,* Averbeck and Hample's (2008) findings suggest that the inclination to use irony can be associated with aggression-related communicative tendencies (such as verbal aggressiveness).

As will be hypothesized, both the detection and the use of irony base on the receiver's ability-dependent cognitive processing and trait-dependent preferences. That is, cognitive and affective processes involved in irony detection and use may depend on traits and abilities from the realm of established individual differences variables. As a construct that delineates the successful performance on a variety of tasks that are cognitively demanding, general mental ability (intelligence) should be expected to play a role, especially in irony detection. This may be the case because irony detection is assumed to involve a cognitive challenge beyond the demand that the processing of literal speech poses on the individual (such as

when engaging in a cognitive search for the antecedent state of affairs an ironic utterance refers to and, ultimately, for the intended meaning of an ironic utterance).

As to the role of *personality*, it can be assumed that—referring to the conceptual definition of personality traits used by Wilt and Revelle (2015)—the emotions, mental processes, behaviors, and motives involved in irony detection and use may interlink with the stable and enduring affective, cognitive, behavioral, and desire-related patterns that are summarized by personality traits. Hence, it is expected that interindividual variance in irony detection and use can be explained (to a certain degree) by personality and ability, as irony behaviors can be seen as "arena" in which traits and abilities manifest themselves in behavior. Irony use as a social behavior involving the experience and expression of emotions may be a designated behavioral domain in which certain personality traits express themselves. Likewise, the successful detection of irony may be hypothesized to base on the receiver's readiness to process emotionally laden information and to recognize affective states in others; and presumably, this readiness is also a function of certain personality traits.

As irony overlaps with humor, traits that delineate tendencies to understand humor, to appreciate humor, and to perform humor may be especially relevant. Accordingly, humor-related traits such as trait cheerfulness and trait seriousness (as the temperamental foundation of the sense of humor, Ruch, Köhler, & Van Thriel, 1998) and especially the sense of humor itself (e.g., in terms of the virtue-related facets *benevolent humor* and *corrective humor*, Ruch & Heintz, 2016a) may be reasoned to link to irony detection and use.

But why should we care about individual differences in irony detection and use? Irony detection is a prerequisite for professional and social activities, such as negotiating, debating, flirting, teasing, or joking. As a means of ridicule, irony can serve as a tool to point to transgressions of social norms. In this case, understanding the irony is a necessity for learning rules and adapting to group standards. As a means of humor, irony can also help to create

group cohesion and strengthen social bonds. In this case, missing the irony means missing the joke and also missing out on a positive social experience. Likewise, individuals who are not apt in using irony cannot take advantage of these benefits of irony. In this case, irony trainings may help to improve individuals' social performance and well-being. In addition, linking interindividual differences in irony detection to known traits and abilities can help us learn more about the phenomenon of irony, which can be viewed as rich in interpersonal and intrapersonal purposes and functions.

However, before assuming an individual differences perspective (i.e., looking at which traits and ability variables may explain interindividual variance in irony detection and use), it may be instructive to take a close look at the phenomenon of irony first. There are several theories that aim to define conditions that distinguish irony from non-irony and that make assumptions about functions of irony and motives for its use. Some of them also make assumptions about mechanisms involved in the detection of irony. After dealing with the definition of irony, the present author will attempt to summarize and integrate the existing theories of irony—which feature partly complementary, and partly redundant ideas—at the end of the respective section (i.e., "Theories of irony").

Definition of irony

The *Merriam-Webster Dictionary defines* irony as "the use of words to express something other than and especially the opposite of the literal meaning". The term stems from the ancient Greek word *eirōnia,* which means "simulated ignorance". The word *eirōnia* in turn was derived from the ancient Greek *eirōn* (to be translated as "dissembler", which, in turn, according to the *Merriam-Webster Dictionary,* denotes someone who hides under a false appearance). The *Oxford English Dictionary* provides a definition that also includes motives for the use of irony: It defines irony as the "expression of one's meaning by using language

that normally signifies the opposite, typically for humorous or emphatic effect; *esp.* (in earlier use) the use of approbatory language to imply condemnation or contempt (cf. sarcasm [...])". And, as one of the alternative meanings according to the *Oxford English Dictionary*, irony simply means: "dissimulation" or "pretense".

This lexical definition is in line with features of the definitions of verbal irony that several theories of irony suggest. Characterized by an overt pretense, verbal irony is distinct from lies and deception in that speakers actually want the listener to see through their pretense.[1] According to Groeben and Scheele (2003), this characteristic is agreed upon on by all definitions contained in the different theories of irony: when we use irony, we utter something different from what we want to say (especially so by using a choice of words denoting *the opposite* of our true appraisal of circumstances); however, we want the listeners to recognize the dissimulation and to understand the intended meaning of what we say nonetheless. Hence, this characteristic will be taken into account when defining irony in PARTs I-III; i.e., all PARTs of this thesis will use the same definition of irony, with more or less the same wording. Characteristically, there is a contrast between the valence of the literal utterance and the valence of speakers' true appraisal of circumstances. This leads to a basic distinction between two basic formal types of irony: mock positive evaluation of negative circumstances (i.e., intended as criticism, "ironic criticism", Dews, Kaplan, & Winner, 1995;

[1] Although irony is conceptually distinguished from lies and deception, it is of course thinkable that individuals use irony deceitfully; i.e., an ironic utterance can be used to convey untruthful information against the better knowledge of the speaker. To illustrate: imagine that you show up for work late and make up an excuse by saying "My train was cancelled due to a technical defect. I guess this is my lucky day!". This would be an example in which the ironic self-pity is a way to make the lie even more convincing. Also, a speaker may use irony in order to act *as if* he or she had a negative attitude towards a given circumstance although this is not his or her true appraisal. To illustrate: imagine that you find yourself in a circle of colleagues from work who eat organic food from adequate animal housing only. You may pretend to adapt to their critical attitude toward the low-budget meat-industry by ironically and cynically joking when saying "These cows are lucky to be released from their miserable lives as soon as they are fat enough!". In case you do not truly adhere to the concern of animal welfare when buying low-budget meat in private, this would be an instance of irony used in the context of covert dissimulation.

"ironic insult", Dews & Winner, 1995; "canonical irony", Kreuz & Link, 2002; "the use of approbatory language to imply condemnation or contempt", the *Oxford English Dictionary*) and mock negative evaluation of positive circumstances (i.e., intended as praise, "ironic compliment", Dews et al., 1995; "noncanonical irony", Kreuz & Link, 2002). An example for ironic criticism would be saying "Nice shot!" when a player tries but fails to score a goal in a sports match (e.g., given that the speaker wants to ridicule his own or someone else's positive expectation that the team would win the match because he or she now sees this expectation fail), whereas saying "Terrible shot!" if a player scores a goal would be an example for an ironic praise (e.g., given that the speaker wants to ridicule his own or someone else's negative expectation that the team would lose the match because he or she now sees this expectation fail). The distinction between "canonical" (i.e., established) and "noncanonical" irony made by Kreuz and Link (2002) denotes the phenomenon that ironic statements are commonly theorized to be more typically *ironic criticism* than *ironic praise* (which in the following will be the term used for mock negative evaluation of positive circumstances).

Schmidt-Hidding (1963) lists irony among other styles that constitute the *comic* (in terms of the aesthetic domain encompassing phenomena that are able to make us laugh or to amuse; cf. Ruch, 1998), such as wit, satire, and humor. According to Schmidt-Hidding (1963), irony (a) aims at creating a mutual sense of superiority toward a third, (b) targets a single situation (as opposed, for example, to *cynicism*, which has the weak world as its object), (c) is characterized by the agent's conceited, superior, relaxed, frequently negative-critical attitude, (d) is courting and including the intelligent while mocking the stupid, (e) works best with a circle of insiders or the informed, (f) is confusing to non-insiders and leaves the judgment remaining up in the air, and (g) is double-faced and characterized by the linguistic peculiarity of saying something differently to how it is meant.

In addition to the definition referred to above, in the present thesis an utterance will be defined as irony (a) if it is purposefully intended as irony by the sender *and* understood as such by the receiver (in terms of a correct positive detection), or (b) if it is purposefully intended as irony by the speaker *and not* understood as irony by the receiver (in terms of a false negative detection of irony). It is important to state that this conceptual definition is simplified (i.e., it invokes the two extremes of the possible outcomes of irony detection), because in the present thesis the detection of an ironic "signal" will not be operationalized as a dichotomous outcome in terms of "detected" or "not detected". Rather, since ironic utterances (and in the operationalization of irony detection: ironic stimuli) conform with "prototypical irony" to varying degrees (cf. Utsumi, 2000), it is assumed that there is uncertainty involved in the detection process (i.e., when listeners are not sure whether an utterance is ironic or not). Accordingly, accounting for this uncertainty and the ambiguous nature of ironic utterances, in the present thesis irony detection will be assessed using a continuous (i.e., rating-scale based) rather than a dichotomous answer format (see PART II).

Correspondingly, an utterance will be defined as non-ironic if no irony is intended by the speaker and none is understood by the receiver (in terms of a correct rejection of irony). What if a receiver detects irony although the sender intends none? In this case, it will be argued that—in terms of purposeful and reasonable communication—there was no ironic "signal" present in the first place and the detection of irony is a false alarm (i.e., the utterance *per se* is non-ironic)[2]

[2] The phenomenon of false alarms in irony detection was mentioned as being worthwhile of study (Kreuz, 2000). Accordingly, Bruntsch, Hofmann, and Ruch (2016, i.e., PART I) hypothesized that the rationale and scoring of future tests of irony detection should account for false negative irony detection. They also hypothesized that individuals with high expressions of the fear of being laughed at (gelotophobia) may be more prone to false alarms in literal compliments (as compared to fear-free individuals) because they have a tendency to anticipate being ridiculed. As a special case of false positive irony detection, non-ironic utterances were found to involuntarily create an ironic meaning if the speaker does not have crucial knowledge possessed by the receiver or the audience (Gibbs, O'Brien, & Doolittle,

Demarcation between verbal irony and situational irony

Although both are referred to as "irony", verbal irony and situational irony are distinct phenomena. According to the *Merriam-Webster Dictionary*, irony also delineates "a situation that is strange or funny because things happen in a way that seems to be the opposite of what you expected". More specifically, the *Merriam-Webster Dictionary* defines that irony, in this situational sense, denotes an "incongruity between the actual result of a sequence of events and the normal or expected result" or "an event or result marked by such incongruity". This is in line with the definition given by the *Oxford English Dictionary* that irony also denominates "[a] state of affairs or an event that seems deliberately contrary to what was or might be expected; an outcome cruelly, humorously, or strangely at odds with assumptions or expectations". According to Lucariello (1994), irony as situational irony is characterized by unexpectedness but also by the exposure of human fragility or the illustration of "the vulnerability of the human condition" (p. 129) by mocking the normal order of things. The *Merriam-Webster Dictionary* lists an example that illustrates Lucariello's (1994) definition of situational irony: "The great *irony* of human intelligence is that the only species on Earth capable of reason, complex-problem solving, long-term planning and consciousness understands so little about the organ that makes it all possible—the brain. —Amanda Bower, *Time*, 20 Aug. 2001". Bower describes the human weakness that we are not able to understand what makes us understanding beings. As one of the categories of situational ironies listed by Lucariello (1994), one also finds a logical "Catch-22" in this example: In order to make humans understanding beings, the brain needs to be more complex than they

1995). In this case, the speaker is honestly misled by a lack of knowledge, while the listener has knowledge of true facts that reverse the circumstances (so a mistakenly sent literal utterance turns into a true received ironic utterance). However, according to the definition that ironic utterances must be purposefully intended as irony, a mistaken utterance based on false or missing knowledge will not be considered as irony here, even if listeners think it is irony.

can understand; even if they were much smarter they would not understand the brain, because in this case the brain presumably would have to be even more complex in order to make humans smarter.

Demarcation between irony and sarcasm

Irony and sarcasm are overlapping but not interchangeably defined phenomena. While the *Oxford English Dictionary* defines sarcasm without the mention of irony as "A sharp, bitter, or cutting expression or remark; a bitter gibe or taunt.", in the *Merriam-Webster Dictionary* sarcasm is defined as "a sharp and often satirical or ironic utterance designed to cut or give pain". Further stressing the role of irony in sarcastic speech, the *Merriam-Webster Dictionary* explicates sarcasm also as "a mode of satirical wit depending for its effect on bitter, caustic, and often ironic language that is usually directed against an individual". It can be observed that, foremost in the Anglo-American literature, many of the studies dealing with irony prefer the term *sarcasm* to the term *irony*. This might be owed to the fact that these studies are using sarcastic irony (for example, in the form *of ironic criticism*) in their stimulus materials. As the present thesis aims to also investigate non-sarcastic forms of irony (i.e., benevolent, playful, or teasing irony), the term sarcasm will not be used unless the reporting of pre-existing studies that used this term makes it necessary to reuse it. As sarcasm does not necessarily involve irony, and also because those irony studies that use the term sarcasm always refer to ironic sarcasm, the present thesis will stick to the term *irony*.

Four domains of irony behavior

The present thesis targets the investigation of irony detection aptitude and the tendency to use irony as two domains of irony behavior. Importantly, these are not the only two domains in irony behavior that have been investigated in the existing literature. In Appendix A1, irony

behavior is described along two orthogonal dichotomous taxonomic aspects: (1) irony reception vs. irony production, and (2) irony ability (i.e., maximal behavior) vs. irony preference (i.e., typical behavior). As Appendix A1 shows, there are four domains of irony behavior resulting from the combination of these two taxonomic aspects. These domains are irony reception ability (i.e., irony detection aptitude), irony production preference (i.e., irony use tendency), irony production ability (i.e., irony creation ability), and irony reception preference (i.e., irony appreciation). While the former two are investigated in the present thesis, the latter two are not. This selection is also owed to the fact that there is only very little previous work on irony reception as a preference (i.e., irony appreciation, e.g., Blouin & McKelvie, 2012) and no studies at all could be found dealing with irony production as an ability (i.e., irony creation ability). Just like irony detection aptitude and irony use tendency (as investigated in the present thesis), irony creation ability and irony appreciation can be expected to vary systematically between individuals. In Blouin and McKelvie's (2012) study, irony appreciation was operationalized by asking individuals to make ratings of "humor", "profoundness", "their interest", and "liking" of a text containing irony (but also metaphor). As it turned out, an overall appreciation score was negatively correlated with formal thinking and positively correlated with postformal thinking in Blouin and McKelvie's (2012) study (see section "*Postformal thinking* and creativity in irony detection").

As a possible operationalization for the assessment of irony creation ability, individuals could be instructed to generate irony in an open-ended answer format. Successful creation of irony subsequently would need to be rated by trained experts. What makes the operationalization of irony creation ability different from the operationalization of irony use tendency (which can also be assessed using open-ended response formats, cf. Dress, Kreuz, Link, & Caucci, 2008) is that irony creation ability measures can be reasoned to necessitate a performance test instruction. While individuals are requested to respond to the stimuli as if

they would encounter them in real life without any mention of the relevance of irony (cf. Dress et al., 2008; Matthews, Hancock, & Dunham, 2006) when measuring irony use tendency, measures of irony creation ability should reveal to test-takers that they are supposed to generate irony in their responses and to make an effort to achieve a high performance (e.g., to generate as many instances of irony as they can within a given time).

What is irony detection aptitude?

In the present thesis, irony detection aptitude is defined as the individual's ability to successfully receive verbal irony. Irony is successfully received when the intended meaning of an ironic utterance is recognized. Irony detection aptitude is expected to vary across individuals to a substantial amount. Interindividual differences in irony detection aptitude are seen as stable and enduring, as individual's irony detection aptitude is expected to manifest in irony detection performance across a variety of situations, stimuli, or settings in a constant fashion. Irony detection aptitude is assumed as a continuum with high scorers and low scorers framing an approximately normal distribution of this ability across the population. As a theoretical characterization, high scorers typically make sense of verbal irony in no time by recognizing its counterfactual nature and intuitively inferring the true meaning of what the ironic speaker wants to say. Low scorers, on the other hand, get puzzled by ironic utterances, as they do not see through the overt dissimulation the ironic speaker engages in and it is not obvious at all to them what the ironic speaker really wants to say. Low scorers' irony detection aptitude deficits necessitate them to ask ironic speakers what they mean with their ironic utterance or else leave them oblivious to the communication goals of ironic speakers. As measures for the assessment of irony detection aptitude performance tests should be used (see section "Assessment of irony detection and irony use").

What is irony use tendency?

In the present thesis, irony use tendency is defined as the individual's preference (i.e., disposition) to use verbal irony as a means of interpersonal communication. When irony is used, speakers engage in an evident dissimulation when inversing the valence of their true appraisal in the verbatim utterance (i.e., especially by using a choice of words denoting *the opposite* of their true appraisal of circumstances). Interindividual differences in irony use tendency are seen as stable and enduring, as individual's irony use tendency is expected to manifest in irony use across a variety of situations, stimuli, or settings in a constant fashion. Irony use tendency is assumed as a continuum with high scorers and low scorers framing an approximately normal distribution of this preference across the population. As a theoretical characterization, high scorers in irony use tendency make use of irony more often than low scorers and they have a higher preference to achieve their communication goals by using irony rather than literal language (as compared to low scorers). As a necessary condition for *irony use tendency* to manifest itself in irony use, it can be reasoned that high scorers need to have the requisite skill to successfully generate irony in terms of *irony creation ability*. Accordingly, an interaction between the two domains of irony behavior can be assumed, with irony use tendency having a greater impact on irony use among those individuals who are able to successfully generate irony than among those who are not.

Theories of irony

Before assuming an individual differences perspective (i.e., looking at traits and ability variables that may explain interindividual variance in irony detection and use), it may be instructive to take a close look at the phenomenon of irony first. What is irony, why is it used, and which factors play a role when it is detected? The prominent theories of irony will be described in the following sections in order to provide readers with a deeper understanding of

the phenomenon of irony. The theories of irony will be listed in chronological order of their first appearance in the literature. Some of the theories were tested experimentally and are therefore supported by empirical evidence. However, constrained by the aim to provide a concise introduction relevant to the specific topic of the present thesis, those empirical studies that aimed to test the specific assumptions made by the different theories will not be reported here.

What do they attempt to explain?

The theories described in the remainder of this section deal with the phenomenon of irony by attempting to explain (a) which general basic preconditions for natural communication enable ironic communication (i.e., Grice, 1975), (b) why irony is preferred to non-ironic language (Sperber & Wilson, 1981), (c) several features of irony such as the asymmetry of affect, i.e., why ironic criticism is more prevalent than ironic praise (Clark & Gerrig, 1984), (d) which general preconditions must be fulfilled for irony to be understood (i.e., an allusion to an antecedent, Kreuz & Glucksberg, 1989), or attempting to (e) integrate several features of different pre-existing theories (Kumon-Nakamura, Glucksberg, & Brown, 1995), or (f) define the general laws that help to distinguish irony from non-irony (Utsumi, 2000).

However, a theory aiming to explain genuinely psychological aspects of irony, such as emotional antecedents and consequences of irony, could not be found—not to mention the absence of a theory aiming to describe and to explain interindividual differences in different domains of irony behavior. Still, the pre-existing theories have certain implications for the psychology of irony in terms of the general laws in which irony behavior occurs and also for the aim to assume an individual differences perspective on irony behavior. These implications will be discussed at the end of the current section (i.e., “Conclusion on theories of irony”).

Grice's (1975) theory of conversational implicature and the cooperative principle

In his theory of *conversational implicature*, Grice (1975) posits that natural talk exchanges follow a *cooperative principle*, meaning that every participant in a conversation is engaged in the pursuit of a purpose, dedicated to successful and efficient communication. It must be noted that this theory is not specific for irony, but instead aims to define basic preconditions for natural communication in general. However, Grice (1975) exemplifies how certain assumptions he makes hold true in the case of irony. Grice (1975) defines a set of four maxims that shape natural communication. These maxims help listeners to understand what other people want to say—also in the case of irony. The four maxims are: quantity, quality, relation, and manner. It is noteworthy that these maxims are not normative in terms of advisable guidelines. Rather, they are naturally occurring and observable in language in terms of an inherent principle of communication. In other words, Grice does not lecture how us on we can communicate more successfully and efficiently. Instead, he posits that we are used to (and inevitably engaged in) these rules in order to make sense of others' communicative contributions, and to make ourselves understood. The four maxims are characterized as follows: (1) Contributions should be made as informative as is required (for the current purposes of the exchange), but not more informative than is required (maxim of quantity), (2) contributions are expected to be genuine and not false (maxim of quality), (3) contributions are expected to be appropriate for immediate needs at each stage of the transaction (maxim of relation), and (4) conversation partners are expected to disclose clearly what contribution they are making, and to execute their performance with reasonable dispatch (maxim of manner). As irony typically involves counterfactual utterances or statements that do not reflect speakers' true appraisal of circumstances, these utterances can be characterized as insincere. Accordingly, next to metaphor and hyperbole, Grice (1975) lists *irony* as a typical

category of speech acts that flout the maxim of quality. Grice (1975) provides the following example of irony, along with an explanation: "*X, with whom A has been on close terms until now, has betrayed a secret of A's to a business rival. A and his audience both know this. A says 'X is a fine friend'. (Gloss: It is perfectly obvious to A and his audience that what A has said or has made as if to say is something he does not believe, and the audience knows that A knows that this is obvious to the audience. So, unless A's utterance is entirely pointless, A must be trying to get across some other proposition than the one he purports to be putting forward. This must be some obviously related proposition; the most obviously related proposition is the contradictory of the one he purports to be putting forward.)*" (p. 53)

Considering this explanation, how is irony detected according to the theory of conversational implicature and the cooperative principle? In the case of irony, conversational implicature can be described as the process with which the listener recognizes that the ironic speaker ostensibly flouts the maxim of quality by saying something false. Rather than expecting the speaker's contribution to be purposeless, the listener anticipates that the speaker is still committed to the cooperative principle, and subsequently the listener engages in a search for the true meaning of the speaker's utterance. By assuming that the speaker uses irony, the listener finds a sincere contribution on the level of the implicated meaning. In other words, irony is detected if it is the most plausible explanation for untruthful communicative contributions. Because of the cooperative principle, speakers have the option to flout the maxim of quality on the level of what is uttered, and can rely on listeners to make an effort to search for the true meaning on the level of the implication, where the maxim of quality still holds true.

The "echoic mention" theory by Sperber and Wilson (1981)

Sperber and Wilson (1981) define ironic utterances as utterances in which speakers are "literally saying one thing and figuratively meaning the opposite" (p. 295). In their theory of *echoic mention,* they address a basic question: If ironic utterances can be alternatively expressed by their literal opposite, why do we use irony? Or as they put it: "why a speaker should prefer the ironical utterance *What lovely weather* to its literal counterpart *What awful weather* which, on this analysis, means exactly the same thing" (Sperber & Wilson, 1981, p. 295). According to the theory of echoic mention, it is characteristic that, with an ironic utterance, the speaker mentions something that was previously said (i.e., the *proposition*) in order to "make clear that he rejects it as ludicrously false, inappropriate, or irrelevant" (Sperber & Wilson, 1981, p. 308). Sperber and Wilson (1981) point out that Grice's (1975) theory of conversational implicature falls short of accounting for instances of irony for which the criterion of truthfulness does not apply, such as ironical questions, ironical understatements, and ironical references to the inappropriateness or irrelevance of an utterance rather than to the fact that it is false. In other words, Sperber and Wilson (1981) posit that untruthfulness is not a necessary condition for irony; nor is it a sufficient condition, as not every instance of untruthfulness can be considered as irony. In their view, what distinguishes irony from untruthful non-irony is precisely the fact that ironic utterances are cases of echoic mention. It is noteworthy that the proposition, which is echoed in the ironic utterance, does not need to be an explicitly spoken remark. Rather, even if there is no prior utterance that possibly can be echoed, irony can mention unspoken propositions. It may be instructive to look at Sperber and Wilson's (1981) elaboration on the ironic utterance given as an example earlier: "*What lovely weather. Suppose that, as we were deciding to set off on our walk, someone told us that the weather was going to be lovely. It is quite clear that […][What lovely weather] is an ironical echo of this remark. Or suppose we have spent a rainy winter talking about the walks we will have in the summer sun. The echoic quality of*

[…][What lovely weather], though its source is more distant, is nonetheless clear. Even when there is no prior utterance some vague echoing is still involved. One normally sets off for a walk in the hope or expectation of good weather: What lovely weather may simply echo these earlier high hopes." (p. 310).

Sperber and Wilson (1981) provide a list of aspects of irony that can partly be explained by their theory. Three of them are taken up by a later theory (i.e., the pretense theory of irony by Clark and Gerrig, 1984), which is why they will be described in the remainder of the current section. Firstly, Sperber and Wilson (1981) note that there is an asymmetry between two basic forms of irony: (1) ironically saying something positive about something negative, and (2) ironically saying something negative about something positive. According to their account, the first category is used more frequently than the second one. In other words, they state that mock positive evaluations of negative circumstances (which oftentimes are labeled "ironic criticism") are more likely to occur in language use than mock negative evaluations of positive circumstances (oftentimes labeled "ironic praise"). The explanation given by Sperber and Wilson (1981) can be put as follows: standards or rules of behavior that generate positive expectations (which, when violated, can be addressed by ironic criticism) are more available than those standards or rules of behavior that generate expectations of negative outcomes (which, when violated, can be addressed by ironic praise). Hence, for ironic praise "[i]n the face of a perfect reality, there must be past doubts or fears to echo if the mention of a critical judgment is to count as ironical" (Sperber & Wilson, 1981, p. 312).

Secondly, Sperber and Wilson (1981) discuss that irony characteristically is aimed at a particular target or victim. On the one hand, ironic utterances are argued to carry critical overtones targeted at the originators of the utterances or opinions being echoed (as the ironic speaker rejects them as "ludicrously false, inappropriate, or irrelevant" (Sperber & Wilson,

1981, p. 308). On the other hand, irony can be victimizing if the addressee of the ironic utterance publicly fails to detect its figurative meaning (i.e., the meaning intended by the ironic speaker) and gets excluded from the "conspiracy" (Sperber & Wilson, 1981, p. 313) between the speaker and third parties. As described in the section "Definition of irony", this aspect was also mentioned by Schmidt-Hidding (1963), who defined that irony works best with *the circle of insiders* or *the informed*.

Thirdly, Sperber and Wilson (1981) acknowledge the existence of an ironic (synonym: ironical) tone of voice. They point out that the ironic tone of voice "is merely one of the variety of tones (doubtful, approving, contemptuous, and so on) that the speaker may use to indicate his attitude to the utterance or opinion mentioned" (p. 311).

Sperber and Wilson (1981) point out that their theory provides an advancement over Grice's (1975) theory but leaves some problems unaddressed, as "a mention may be more or less ironical, with many intermediary and complex shades between stereotypical cases of irony and other kinds of echoic mention." (p. 315). This is an important notion, as it demonstrates that the conditions constituting irony in this theory (and also in the other theories of irony, e.g., the condition of insincerity) are not necessarily specific to irony (i.e., echoic mention and insincerity can also be used non-ironically). Furthermore, this notion is important because the conditions assumed for irony may occur in ambiguous or transitional varieties rather than in a dichotomous fashion. Addressing this issue, the *implicit display theory of irony* (Utsumi, 2000) accounts for the aspect of ambiguity, as in one of its axioms it includes the degree to which an ironic utterance deviates from prototypical irony; see section "The implicit display theory of verbal irony by Utsumi (2000)". Nonetheless, due to the richness of observations, ideas, and discussions offered, Sperber and Wilson's (1981) essay on the echoic mention theory inspired several subsequent theories of irony.

The pretense theory of irony by Clark and Gerrig (1984)

According to the pretense theory by Clark and Gerrig (1984), irony is characterized by speakers pretending to assume the role of a person who imprudently believes in the literal meaning of the ironic utterance. The speaker, however, expects the listeners to see through this dissimulation and thereby discover that the speaker has an attitude toward the person he or she is impersonating. The pretense theory aims to provide explanations for several features of irony mentioned by Sperber and Wilson (1981), namely (1) the asymmetry of affect, (2) victims of irony, and (3) the ironic tone of voice. Clark and Gerrig (1984) explain the asymmetry between ironic criticism and ironic praise (which they call *asymmetry of affect*) by defining irony as a pretense in which the speaker assumes the role of persons who see the world according to norms of success and excellence. By pretending to "view the world through rose-colored glasses" (Clark & Gerrig, 1984, p. 122), the ironic speaker ostensibly adopts a mentality of ignorance characterized by positive expectations and hence ironic criticism is more likely to occur than ironic praise.

Furthermore, Clark and Gerrig (1984) claim that the echoic mention theory of irony cannot distinguish between two types of victims of irony. According to Clark and Gerrig (1984) the first kind of victim, in terms of the pretense theory, is the "unseeing or injudicious person the ironist is pretending to be" (Clark & Gerrig, 1984, p. 122). The second kind of victim is the uncomprehending audience, which is not in the inner circle of listeners who discover the pretense. However, the incremental explanatory value of the pretense theory over the echoic mention theory can be seen as not very noticeable for this aspect.

The ironic tone of voice can be defined as the phenomenon that an ironic utterance is spoken with a different vocal expression and emphasis than its both its literal counterpart and identical utterances without ironic intent. Clark and Gerrig (1984) claim that the pretense theory accounts for this phenomenon, as "[i]n pretense or make-believe, people generally

leave their own voices behind for new ones." (p. 122). The ironic speaker is supposed to assume a voice that is appropriate for the person he or she is pretending to be, while exaggerating and caricaturing this voice.

The *echoic reminder theory of verbal irony* by Kreuz and Glucksberg (1989)

In their *echoic reminder theory of verbal irony*, Kreuz and Glucksberg (1989) take up ideas entailed in Sperber and Wilson's (1981) theory of echoic mention. Kreuz and Glucksberg (1989) state that ironic utterances characteristically have to allude to an antecedent event if they are to be understood. They claim that their theory extends the theory of echoic mention by also accounting for instances of irony in which the antecedent is not explicitly or implicitly mentioned in an echoic utterance. Rather, irony works if it alludes to an antecedent state of affairs. By reminding the listener of an antecedent state of affairs, the ironic utterance exposes a discrepancy between what is and what should be. Kreuz and Glucksberg (1989) claim that this principle of irony can also account for the phenomenon of asymmetry between ironic criticism and ironic praise. In line with Sperber and Wilson's (1981) reasoning, Kreuz and Glucksberg (1989) explain that positive statements can more readily be used ironically, because positive norms are naturally more prevailing than negative expectations. Accordingly, an ironic positive evaluation of negative circumstances (i.e., ironic criticism) can use these ubiquitous positive norms as antecedents. Ironic negative evaluations of positive circumstances (i.e., ironic praise), on the other hand, cannot make use of such consented norms and need an explicit antecedent if they are to be understood as irony. Typically, the antecedent in ironic praise leads to negative expectations, which—when violated—can be reminded of by an ironic utterance in order to hint at the discrepancy between what is and what should be, on grounds of the antecedent state of affairs.

Allusional pretense theory by Kumon-Nakamura et al. (1995)

In their *allusional pretense theory of discourse irony*, Kumon-Nakamura et al. (1995) integrate and extend Grice's theory of conversational implicature, the theory of echoic mention (i.e., echoing something that was previously said or echoing conventional, foremost positive, expectations; Sperber & Wilson, 1981), and echoic reminder (i.e., reminding the listener of an antecedent event, shared norms, or mutual expectations in order to hint at a discrepancy between what is and what should be; Kreuz & Glucksberg, 1989) to arrive at a more comprehensive theory of irony. Supporting Sperber and Wilson's (1981) view, Kumon-Nakamura et al. (1995) point out that Grice's theory is restricted to utterances that are declarative assertions, for which the criterion of truth applies, and they note that Grice's theory does not include ironic utterances other than declarative assertions that are reversible into their opposite. According to Kumon-Nakamura et al. (1995), irony also involves counterfactual assertion for which the criterion of truth does not apply, but for which the criterion of *sincerity* applies, such as compliments, questions, and offers. To include those categories of irony, they use the more comprehensive term "pragmatic insincerity". The term "pragmatic insincerity" refers to Grice's theory, which—because it claims that conversation partners are inevitably engaged in reasonable and practical communication—is often called "pragmatic theory". Pragmatic insincerity does not only involve uttering the opposite of what we mean, but also other forms of insincerity. For example, a question cannot be true or false, but it can sincerely demand an answer (e.g., if you say to a friend who while he lifts a huge box to help you move into a new flat: "Are you sure you can carry this all the way upstairs by yourself?") or be uttered insincerely in form of an ironic remark (e.g., if you ask this same question when your lazy friend is carrying a very small item). Furthermore, Kumon-Nakamura et al. (1995) agree that the echoic irony is one option to allude to unfulfilled

expectations, but they posit that there are many more ways of calling "the listener's attention to some expectation that has been violated in some way" (p. 5).

Accordingly, Kumon-Nakamura et al. (1995) posit that ironic speech acts are characterized by two conditions: firstly, they allude to expectations that have not been met. Secondly, ironic utterances are defined by pragmatic insincerity because "they violate one or more of the felicity conditions for well-formed speech acts." (p. 5). As a possible function of irony, Kumon-Nakamura et al. (1995) suggest that it helps to implicitly communicate expectations and an "attitude toward the expectation and the discrepancy between the expectation and reality" (p. 21), which characteristically (but not exclusively) is a negative attitude.

The implicit display theory of verbal irony by Utsumi (2000)

With his *implicit display theory of verbal irony* Utsumi (2000) aims to explain how listeners distinguish irony from non-irony, that is, he defines criteria that are necessary for irony detection. In his essay, Utsumi (2000) claims that preexisting theories fall short of sufficiently explaining how listeners judge whether an utterance is ironic or not. The implicit display theory makes three central claims. Firstly, for a listener to detect irony, an ironic environment is required. The ironic environment consists of a speaker's failed expectation and the speaker's negative emotional attitude (e.g., disappointment, anger) toward this failed expectation. That is, the listener needs to infer that the discourse situation is a setting that motivates verbal irony. Secondly, the ironic utterance must be recognized as an implicit display of the ironic environment. That is, the listener must realize that the ironic utterance refers to a situation that motivates irony. Implicit display is achieved by alluding to the speaker's expectation, includes pragmatic insincerity by intentionally violating one of the pragmatic principles (cf. Grice, 1975; see section "Grice's (1975) theory of conversational

implicature and the cooperative principle"), and expresses indirectly the speaker's negative emotional attitude towards his or her failed expectation. Thirdly, irony is distinguished from non-irony by assessing the degree to which a given utterance resembles prototypical irony. That is, since not every ironic utterance has all the properties required for the implicit display of an ironic environment—or unambiguously fulfills the constituting criteria—the listener detects irony by assessing the similarity between a given utterance and a prototype of irony. The similarity of an utterance with the ironic prototype is judged by assessing to what degree the utterance (1) is coherently related to the speaker's expectation, (2) violates pragmatic principles, and (3) indirectly expresses a negative attitude through the use of ironic cues (e.g., hyperbole, prosody, nonverbal behavior).

Conclusion on theories of irony

One of the essential features of irony that the discussed theories of irony agree on is that it involves an overt pretense (cf. Groeben & Scheele, 2003). The speaker utters something denoting the opposite of their true appraisal of circumstances; importantly, this dissimulation is supposed to be recognized by the listener, so the speaker wants to be understood with their ironically implied meaning.

The described theories are important for the aims of the present thesis in several ways: firstly, at least one of them suggests certain general psychological laws in which irony occurs, such as emotional antecedents and consequences of irony. To illustrate: Sperber and Wilson (1981) discuss that irony characteristically is aimed at a particular victim and argue that ironic utterances carry critical overtones targeted at the originators of the utterances or opinions being echoed (as the ironic speaker rejects them as "ludicrously false, inappropriate, or irrelevant" (Sperber & Wilson, 1981, p. 308). It may be suggested to derive certain predictions as to the emotional antecedents and consequences from this notion. That is,

ridiculing criticism (a) can be seen as typically *motivated* by certain emotions (such as anger and contempt, see the definition of irony in the *Oxford English Dictionary* as "the use of approbatory language to imply condemnation or contempt"), and (b) can be reasoned to represent an antecedent typically *eliciting* certain emotions (such as shame, embarrassment, and guilt, Tangney, Miller, Flicker, & Barlow, 1996). These assumptions can be seen as substantiated by the findings discussed in the sections "Why do we use irony*?*" and "Emotional consequences of irony".

Secondly, some of the described theories have implications for individual differences aspects of irony. To illustrate: according to Grice's (1975) theory, speakers have the option to flout the maxim of quality on the level of what is uttered and, because of the cooperative principle, can rely on listeners to make an effort to search for the true meaning on the level of the implication, where the maxim of quality still holds true. This search for the intended meaning can be reasoned to involve cognitive processes that depend on individuals' general mental ability and sense of humor. Accordingly, in PART II it will be argued that ironic praise as a form of irony that can be assumed to be more demanding than ironic criticism when employed in an irony detection task is more strongly associated with intelligence and certain facets of the sense of humor.

Furthermore, according to the pretense theory by Clark and Gerrig (1984), who claim that irony involves a kind of act or role-play, irony use can be linked to the inclination to engage in as-if behaviors in terms of the histrionic self-presentation style (cf. Renner, Enz, Friedel, Merzbacher, & Laux, 2008). For example, according to Clark and Gerrig (1984), the ironic speaker is supposed to assume a voice that is appropriate for the person they are pretending to be, while exaggerating and caricaturing this voice. This description can be linked to Renner et al.'s (2008) definition of as-if behaviors that are described to "involve more subtle forms of communication and may include the use of metaphors, irony or

humorous word games, grimaces, rolling one's eyes or modulating one's voice etc." (p. 1304).

Linking irony to traits associated with ridicule, Kumon-Nakamura et al. (1995) suggest that, as a possible function of irony, it helps to implicitly communicate expectations that, characteristically but not exclusively, are a negative attitude (see section "Why do we use irony?" of the present thesis). Accordingly, in PART III, those traits that delineate laughing-at, aggressive, or critical humorous behavior, such as *katagelasticism* (Ruch & Proyer, 2009) or *the aggressive humor style* (Martin, Puhlik-Doris, Larsen, Gray, & Weir, 2003) were examined.

The theories of irony mentioned thus far can also be used as guidelines for generating ironic stimuli for the assessment of irony detection and use. As Kreuz and Glucksberg (1989) point out, irony typically refers to an antecedent event or state of affairs that leads to an expectation, which—when violated—can be reminded of by an ironic utterance in order to hint at the discrepancy between what is and what should be, on the grounds of the antecedent state of affairs. For example, in order to facilitate the use of irony in an experimental task, the stimuli should provide an antecedent event generating an expectation. Furthermore, this expectation needs to be violated in order to give the speaker (or, ultimately, the participant) the chance to critically remind the listener of the discrepancy between expectancy and reality, for example, by echoic mention. Accordingly, the stimuli used in the forced-choice irony use measure employed in PART III were adapted by including the violation of an expectancy, such as when the addressee of the ironic remark does not stick to their own rules.

With regard to the assessment of irony detection, Utsumi's (2000) theory can be seen as particularly relevant for the present thesis, as the measurement approach chosen in the studies conducted for PART II involves ambiguous stimuli. Utsumi's (2000) theory accounts for varying degrees of certainty with which irony is being detected. The notion that there may

be imperfect similarity between an utterance and prototypical irony can be seen as supporting the rationale of the irony detection measure developed in PART II of the present thesis, which involves stimuli that, due to their ambiguous nature, may not be classified as irony with a perfect consensus or certainty.

Why do we use irony?

One aim of the present thesis is to test the possibility that (a) interindividual differences in irony use are habitual (i.e., trait-like) and, if so, (b) that these differences are associated with personality traits. In order to build hypotheses about which traits matter for the prediction of the activities and processes denoted by these two domains of irony behavior, it may be worthwhile to take a look at the motives that speakers have when using irony, the functions irony can have in social interaction, the benefits of irony over literal communication, and which communication goals are typically pursued with irony. The motives, functions, benefits, and goals that go along with each other in irony use may be entangled with the experiential and behavioral tendencies typical of high scorers in certain personality traits.

As Kreuz, Long, and Church (1991) reason, if irony is used despite the risk that the listener might misunderstand the speaker by taking the irony literally, irony must have benefits over literal communication. They point out that irony makes it possible for the speaker to fulfill communication goals that would be more difficult, if not impossible, to fulfill literally. In an attempt to explain why ironic utterances were previously found to be memorized better than their literal counterparts, Kreuz et al. (1991) collected ratings of various communication goals for the scenarios they used in their recall study. Although they found that their participants recalled scenarios that ended with an ironic utterance only marginally better than scenarios with a literal final remark in their main analysis, they report that irony involved certain communication goals to a higher degree in their study, as more

participants indicated that they thought that "to be funny or to be witty", "to play or to be silly", or "to mock" was a communication goal fulfilled foremost by the ironic statements (as compared to the literal statements). Kreuz et al. (1991) suggest that as a benefit of irony, speakers can achieve their communication goals highly effectively. Furthermore, they conclude that ironic statements are high in pragmatic involvement, for example, to mock, to insult, or to be mean or funny. These findings are in line with the results reported by Roberts and Kreuz (1994), who asked their participants to generate responses to the question why they think irony is used (among other forms of figurative language, such as metaphor). 64% of the participants generated answers falling into the category "To be humorous", and 94% even responded in a way that was classified as "To show negative emotion". Between these two categories, irony was the highest-ranking form of figurative language, that is, for no other form did participants generate more responses in the categories "To be humorous" and "To show negative emotion".

So why do we use verbal irony? In light of the two studies described in the current section, this question could be answered on two levels: (a) because it represents a way of achieving one's communication goals efficiently, and (b) because we want to mock, to show negative emotion, to be funny or witty, to be humorous, to play, or to be silly. The view that irony can achieve more communication goals than the respective literal translation of irony is supported, for example, by Oomen (1983). According to Oomen (1983), ironic utterances are preferred over non-ironic phrasing by speakers when its surplus over literal communication is desired and the speaker intends to express disappointment, blame, or criticism toward something else than the subject of the ironic utterance.

Concerning the question of which goals are typically achieved by irony, the lists by Kreuz et al. (1991) and Roberts and Kreuz (1994) could still be extended. As irony has been mentioned as a means of playful teasing in close interpersonal relationships, we could add

bonding to the list of functions of irony. As Norrick (1994) points out, "mocking and sarcasm are by definition aggressive in attacking an interlocutor and violating the norms of politeness, albeit within a play frame" (p. 429). By breaking with the norms of politeness and respect when using mocking irony, ironic speakers can be argued to convey a bonding metamessage in terms of ironic teasing (Keltner, Capps, Kring, Young, & Heerey, 2001). This is because they demonstrate that they think that the relationship with the communication partner is beyond the "politeness stage", the stage during which everyone is trying to be respectful and friendly in order not to offend each other. Showing that the relationship is close enough not to require formalities, ironic teasing hence may strengthen the bond between friends. Close friends would criticize each other for their transgressions but would prefer to do so in a benevolent or playful frame. Using irony as a jocular way of conveying a critical attitude (e.g., Garmendia, 2014) may bond individuals because it signals the speaker's appraisal that the relationship is strong enough to stand teasing. Or, as Norrick (1994) puts it "sarcasm and mocking can express both aggression and solidarity - aggression in the message, attacking others for their foibles and errors, and solidarity in the metamessage, including others in a playful relationship with increased involvement" (p. 423). Substantiating this notion, both teasing and irony were reported to be more prevalent among friends (or other close individuals) than among strangers (cf. Keltner et al., 2001; Pexman & Zvaigzne, 2004).

Psychological approaches to humor

Although the phenomena of humor and irony are frequently mentioned together in the existing literature, and their relationship has been purposefully discussed (e.g., Dynel, 2014; Hirsch, 2011; Littman & Mey, 1991), there is no consensus on how humor and irony interrelate. Certainly, not every instance of irony can be seen as an instance of humor (and vice versa). However, the phenomena of humor and irony can be seen as naturally

overlapping to a certain degree (see Bruntsch et al., 2016, i.e., PART I of the present thesis, for a more detailed definition and discussion). Likewise, many of the research questions and conceptual approaches used in the investigation of the psychology of *humor* may also apply to the investigation of the psychology of *irony*.

Humor has been studied across different scientific disciplines, but, importantly, the term *humor* did not consistently denote the same phenomena in studies typically conducted in the different fields of humor research. Oftentimes, humor scholars targeted the investigation of humorous products, for example when the "mechanics" of jokes or the semantic and pragmatic characteristics of humorous texts were analyzed in linguistic studies (e.g., Attardo, 2001a; Raskin, 1985). As a second prominent approach, the study of the characteristics of individuals (but also natural groups or cultures), performing or consuming humor in certain qualities or quantities, can be named (cf. Ruch, 2004). Psychological studies of humor typically follow this second approach.

So, what are typical psychological approaches to humor? As a first step, this question could be confined to account for the focus of interest in psychological studies of humor: What kinds of humor *behavior* are there? According to Ruch's (2008) description, humor behavior refers to a person's activities but also internal processes (such as thinking or emotional feeling) involved in the psychology of humor. Based on an extensive literature review, Ruch (2008) discusses that humor behavior among further aspects involves (a) the perception that something is *funny*, (b) typical responses, like laughter and foremost smiling, (c) cognitive processes, such as the integration of contradiction or the resolution of incongruity, (d) motivational antecedents (such as repressed sexuality or aggression but also feelings of interpersonal superiority), (e) mood and other states (such as playfulness or low seriousness), (f) personality traits delineating humor behavior (in the sense of what is called the *sense of humor* in everyday language). Among the latter, Ruch (2008) distinguishes

between ability (in the sense of maximal behavior) and style (i.e., typical behavior in the sense of habitual preferences). Ruch (2008) conceptualizes that the *sense of humor* has different domains when disentangling it as (a) humor ability (such as in the sense of *wit*), (b) humor as a an habitual preference (such as in the sense of playfulness or different styles of humor use), (c) humor as a character strength (in terms of a disposition to virtuous humor behavior), and (d) humor as an aesthetic perception (i.e., individuals' "taste" in humor). Among the research questions that Ruch (2008) lists for the investigation of humor behavior, some also are applicable (or already have been applied) in the investigation of irony, such as humor pathologies (i.e., humorlessness), development of humor over the life span, factors that support or impede humor, intervention programs, cross-national and cross-cultural perspectives, heritability, and evolution (i.e., phylogenetic development). Hence, if we leave aside the notion for a moment that irony and humor may overlap to a certain degree, the investigation of humor can be seen as a rich source of paradigms for the investigation of irony.

Aggression and humor as motives for irony use

Although there are other functions of irony mentioned in the literature, such as venting frustration (Ducharme, 1994) or demonstrating power by emotional self-control (Dews et al., 1995), two of the most pervasively reported functions of irony can be seen in its overlap with the phenomena of aggression and humor. Although there is a debate about whether irony increases the condemnation of a criticism (cf. Colston, 1997) or whether it dilutes the condemnation and hence reduces the threat to the speaker's face, making further conflict less likely (cf. Dews et al., 1995; Jorgensen, 1996), irony can be noted as being used in aggressive communication. This assumption is supported by the findings of a correlational study by Averbeck and Hample (2008) that jointly investigated irony use tendency and

communication-related traits relevant to aggression (i.e., verbal aggressiveness and argumentativeness). As their study showed, irony use was positively correlated with self-reported individual aggressive communication tendencies.

Considering that irony was found to be perceived as being used when speakers intent to mock, to insult, to be mean, or to be funny (cf. Kreuz et al., 1991), irony may also be relevant for the overlap between aggression and humor, which can be seen in the phenomenon of disparagement humor (Zillmann, 1983; cf. PART I of the present thesis). As irony can be argued to be aggressive and funny at the same time, humor that victimizes its targets by means of mockery, debasement, and ridicule may employ irony in order to achieve its humorous and caustic effect. As argued in PART I of the present thesis, the debasing or ostracizing effect of disparaging irony may be found in the immediate target of ridicule, whereas the humorous effect may be found in bystanders witnessing disparagement humor. It could be argued that derisive laughter or amusement elicited in a group of bystanders aggravates the shaming effect of disparagement humor, possibly making it a suitable tool for punishing the transgression of social norms or misbehavior in a group (cf. Shott, 1979). In line with this assumption, Norrick (1994) points out that sarcasm helps to apprise the recipient of social norms. Moreover, Norrick (1994), once again bridging irony and teasing, sees teasing as a means of enforcing norms. This view is supported by Keltner et al. (2001), who came to the conclusion that "[i]ndividuals often tease others who have violated social norms" (p. 237). Teasing can be seen as running on a continuum: as playful form of provocation, it may be a means of bonding, whereas on the other side of the spectrum it may resemble "biting" ridicule (Boxer & Cortés-Conde, 1997; Keltner et al., 2001). One may argue that ironic teasing runs along the same spectrum between playful, benevolent provocation and disparaging ridicule. Accordingly, irony may be one tool for humor within humor's entire spectrum that spans between the poles of compassionate humor on the one

hand and disparagement humor on the other hand. Following this logic, one must expect that there are humor phenomena in between these morally valued extremes, which can be found true, for example, in the case of corrective humor, which is typically used for shaming corrupt individuals into betterment (Ruch & Heintz, 2016a).

Conclusion

As previously suggested, individuals view irony to be used both in order to achieve a funny or amusing effect on the one hand and to mock or show negative emotions on the other. Since irony may be involved in both, the phenomena of humor and aggression, irony may have a special role when it comes to the disparaging part of the humor spectrum. Therefore, in addition to personality dimensions relating to aggression (i.e., psychoticism), the present thesis will take on humor-related traits that not only delineate playful compassionate aspects of the sense of humor (such as *benevolent humor*; cf. Ruch & Heintz, 2016a), but also those traits that go along with a preference for laughing-at humor, such as katagelasticism (Ruch & Proyer, 2009) and the *aggressive humor style* (Martin et al., 2003), or the tendency to use satiric humor with virtuous motives (i.e., *corrective humor*; Ruch & Heintz, 2016a).

Emotional consequences of irony

In the present thesis it is argued that the successful detection of irony may base on the receiver's readiness to process emotionally laden information (and that to a certain degree this readiness is a function of certain personality traits). Therefore, it can be seen as necessary not only to discuss the findings on the motives of irony use (see previous section) but also literature on the emotions typically elicited by irony. There is a considerable body of research investigating how individuals perceive (a) attributes of ironic utterances (e.g., humorousness, politeness), (b) the speaker's motivation to use irony (e.g., status elevation, anger), and (c) assumed effects on the addressee (e.g., how they think the addressee would feel as a result of

a given ironic remark). Importantly, most of these studies make their participants judge these aspects from the perspective of the observer (e.g., Bowes & Katz, 2011; Colston, 1997; Dews et al., 1995; Matthews et al., 2006; Pexman & Olineck, 2002). There are only very few studies targeting emotional reactions to ironic stimuli directly in the addressee of ironic remarks. In one of those studies, Leggitt and Gibbs (2000) explored emotional responses to ironic and non-ironic stimuli by asking their participants to imagine themselves as actually taking part in each of a list of situations, take the perspective of the addressee of the respective ironic remark, and to imagine how they would feel when the speaker said what he or she said to them. Participants reported their emotional reactions along ratings targeting each of a list of emotions. As it turned out, certain forms of irony (e.g. sarcasm, overstatement, rhetorical questions) elicited a higher degree of (a) "angry, irritated, and mad", (b) "disgusted, turned off, and repulsed", and (c) "scornful, disdainful, and contemptuous" emotions (as one category each; Leggitt & Gibss, 2000, p. 8) than the rest of the given types of speech acts that did not involve verbal irony. These findings go along with those of Akimoto et al. (2014) who report that their participants perceived higher degrees of negative emotion and humor in their ironic and non-ironic stimuli. That is, one and the same stimulus stories elicited more perceived negative emotion and more perceived humor (these categories are not explained in any more detail in Akimoto et al.'s, 2014, report) when they were designed as ironic criticism (*mock praise*, i.e., ironically commenting on a failure as if it was a success, see PART II of the present thesis) than when they were designed as literal praise (i.e., commenting on a success by using a positive literal appraisal). Contradicting these findings, another study found that irony dampened the positive emotional effect of praise: employing electro-physiological measures of facial expressions, Thompson, Mackenzie, Leuthold, & Filik, (2016) found less activity in the corrugator supercilii muscle (i.e., reduced frowning) as an indicator of a reduced negative emotional impact of ironic

praise as compared to literal praise. Thompson et al. (2016) interpret their result as a support for the tinge hypothesis proposed by Dews and Winner (1995). Dews and Winner (1995) assume that the presence of a positive verbatim meaning helps speakers to take the edge off their criticisms in the case of ironic criticism (as compared to literal criticism) and likewise that the presence of a verbatim negative meaning mutes the positive appraisal in the case of ironic praise by effecting a more critical evaluation than literal praise. However, there are more recent studies contradicting these assumptions when suggesting that in ironic criticism condemnation is enhanced rather than diluted (as compared to literal criticism) and that ironic insults are perceived to be more mocking than direct insults (i.e., Colston, 1997; Pexman & Olineck, 2002).

Taken together, the emotions elicited in the addressees of irony can be seen as understudied. The existing literature suggests that the emotions elicited by the ironic stimuli used in the described studies are mainly more negative than responses to non-ironic stimuli (as far as self-reported emotions are concerned). As a general methodological issue, it can be seen as questionable whether irony and non-irony are directly comparable in the mentioned studies (in terms of internal validity). For example, ironic praise and literal praise may not only differ in the degree of irony used in communication but also regarding the communication goals the speaker wants to achieve. This may be the case especially if the metamessages of the irony are taken into account. For example, in ironic teasing (cf. Keltner et al., 2001) speakers may wish to be playful and critical at the same time. Hence, if an instance of teasing ironic praise is compared to an instance of literal praise as to the emotions they elicit, one does not only compare irony to non-irony but also teasing to non-teasing. Furthermore, given that (a) irony serves as a means to achieve various communication goals and (b) the mentioned studies did not sample their stimuli in a representative fashion with

regard to these communication goals, it is not clear to which extent the described findings can be generalized.

Sources of variance in irony detection and use described in the existing literature

The idea of investigating person-specific variables in the context of irony behavior is not new. However, the vast majority of these studies were not looking at normal functioning adults' irony detection and use. Rather, groups known for impaired irony detection performance or impairments abilities assumed to be a prerequisite for irony detection were compared to normal adults. As the following sections will describe, there is a plethora of research reports dealing with underperformance in irony detection among children, individuals with an autism spectrum disorder, and among patients suffering from schizophrenia, posttraumatic brain injury, Parkinson's disease, or Alzheimer's disease. Concerning irony use, some studies have focused on person-specific variables as correlates of irony use, such as culture, gender, or traits relating to aggression.

Ontogenetic aspects of irony detection

The detection of irony has been studied extensively in children (cf. Anglieri & Airenti, 2014). The reason for testing children can be seen in the fact that the cognitive prerequisites for irony detection are assumed to develop at a young age. These prerequisites include verbal ability and, in particular, the ability to infer mental states in others, which is commonly labeled using the term *Theory of Mind*. While *Theory of Mind* is assumed as fully developed in normally functioning adults, with only little interindividual variance (cf. Brüne & Brüne-Cohrs, 2006), studying irony detection in children in different age groups is a paradigm that provides variance in *Theory of Mind*. Commonly, deficits in irony detection among younger vs. older children or adults are explained by the inability to infer others' knowledge and

intentions and to take on others' mental perspectives. In order to realize that a speaker utters something different from what he or she really means, the listener has to represent the speaker's mental state. As successful irony detection is dependent on realizing the speaker's true knowledge and intentions, the developing *Theory of Mind* in young children is seen as one reason for the finding that irony detection rates improve steadily as individuals grow older. For example, Anglieri and Airenti (2014) found increasing irony detection rates between the ages of three and six-and-a-half years. Elsewhere, children were found to understand the non-literal meaning of irony by the age of five to six years, but fall short of understanding the pragmatic functions of irony (i.e., ratings of how mean the speaker was) until the age of nine to ten years (i.e., Glenwright & Pexman, 2010). It may be worth discussing, in more detail than is appropriate for the scope of the present thesis, whether the association postulated and found between irony detection and *Theory of Mind* has to become evident quite naturally when studying children in an age-heterogeneous sample due to the confounding effect of general cognitive development. For example, Anglieri and Airenti (2014) report moderate correlations between irony detection and theory of mind (as measured with three different pre-existing tasks). However, both irony detection and *Theory of Mind* are correlated considerably higher with age in this study. This possible limitation of the developmental paradigm can be seen as support for the relevance of the aim of the present thesis to focus on adults' irony detection and use.

Clinical groups known for *Theory of Mind* deficits

Based on the existing literature, it is hard to distinguish what came first: (a) the observation that certain clinical groups have an impaired aptitude to detect verbal irony, (b) the observation that these same groups have deficits in inferring others' mental states and emotions with regard to the abilities denoted by the term *Theory of Mind*, or (c) the

assumption that *Theory of Mind* is a prerequisite for irony detection. In the literature, there is a plethora of studies comparing irony detection performance between groups of normal functioning individuals on the one hand and groups of individuals known for impaired cognitive functioning in terms of *Theory of Mind* deficits on the other. Among the pathologies that are known to co-occur with reduced *Theory of Mind* deficits, there are patients with schizophrenia, brain damage, and degenerative brain disorders, as well as autism spectrum disorders (cf. Brüne & Brüne-Cohrs, 2006). Tellingly, all of these groups have also been investigated as to their deviation from normal functioning individuals' irony detection performance levels. For example, patients suffering from right hemisphere brain damage were demonstrated to be characterized by lower irony detection performance as compared to the performance found in healthy controls (e.g., Winner, Brownell, Happé, Blum, & Pincus, 1998). Mitchley, Barber, Gray, Brooks, and Livingston (1998) showed that, on average, schizophrenic patients interpreted ironic utterances in a literal way more often than a control group. In line with this finding, individuals with increased risk for psychotic illness in terms of schizotypal personality were found to be less apt at detecting irony (i.e., in terms of reporting that ironic utterances make sense to them; Langdon & Coltheart, 2004). Impaired irony detection performance was also found in individuals with Asperger syndrome and autism disorder (e.g., Happé, 1993; Martin & McDonald, 2004), as well as Alzheimer's disease (e.g., Gaudreau et al., 2013; see Rapp & Wild, 2011, for an overview) and Parkinson's disease (Monetta, Grindrod, & Pell, 2009).

The lowest common denominator of these studies' results can be seen in the finding that deficits in irony detection are associated with deficits in the ability to form representations of others' mental states in terms of *Theory of Mind*, and especially so in deficits when attributing second-order beliefs (i.e., inferring what a person thinks about another person's thoughts; cf. Monetta et al., 2009). Although *Theory of Mind* has been

investigated in infants and clinical groups for many decades, the investigation and assessment of individual differences in *Theory of Mind* among healthy should be viewed as still in its infancy. Furthermore, there is an ongoing debate on a conceptual level about which abilities and processes are actually involved in what is termed *Theory of Mind* (e.g., Schaafsma, Pfaff, Spunt, & Adolphs, 2015).

Conclusion

Despite the above-mentioned shortcomings, this field of clinical irony research is worth mentioning in the present thesis because it illustrates the limitations of this large subset of irony research. These studies demonstrate what makes irony detection possible in normally functioning adults (i.e., *Theory of Mind*). However, the design of these studies can be seen as limited because (a) the informative value about the "laws" in which irony detection deficits occur in the population of normally functioning adults (in which *Theory of Mind* is typically fully developed) can be seen as vague, (b) as internal validity may be reduced (resulting from the fact that impaired *Theory of Mind* is confounded with all kinds of disorder-related variables in clinical studies), and (c) as the method of investigation is costly (because clinical studies are less feasible than investigating irony behavior in the normally functioning population). The individual differences approach has the potential to eventually overcome these limitations (at least to a certain degree) in describing and explaining interindividual differences in irony behavior by linking them to those known personality and ability variables that show stable variance across the population of normally functioning adults.

Culture and gender differences in irony use

Based on the assumptions that (a) irony provides certain benefits in social interaction over literal communication, and (b) these benefits are less desirable for collectivists than individualists, and may also have different desirability for women than for men, Rockwell

and Theriot (2001) investigated the role of gender and individualist vs. collectivist culture in irony use. More specifically, they reasoned that individuals from collectivist cultures (such as the Japanese, Chinese, or Thai), due to their tendency not to express negativity toward other group members, could be expected to use irony less frequently than individuals from individualist cultures (such as the Americans, British, or Germans). Indeed, their collectivist participants used irony less frequently in dyadic face-to-face conversations (i.e., according to self-reports, see below). There was no directed hypothesis regarding gender. As it turned out, women used irony less frequently than men (especially when the women were talking to other women, i.e., in gender-homogeneous dyads). Rockwell and Theriot (2001) explain this finding by women's tendency "to place greater emphasis on maintaining relationships and avoiding injury to conversational partners" (p. 49). The methodological strength of this study can be seen in the fact that face-to-face conversations were investigated. Participants were instructed to deal with 15 questions designed to elicit irony, for example, by requesting participants to utter something negative about positive circumstances (e.g., "Why do you hope you will never win the lottery?"), or something positive about negative circumstances (e.g., "Why do you really like to go to the dentist?"). As a possible limitation of the study, irony use was assessed via self-reports rather than objective ratings: Following the conversations, participants rated how much sarcasm they think they used, as well as the amount of sarcasm used by their conversation partners.

Rockwell and Theriot's (2001) results on the role of gender are in line with the ones found in a study by Colston and Lee (2004). As they expected, the latter authors found that women reported lower endorsement of ironic remarks. In line with Rockwell and Theriot's (2001) reasoning that the benefits of irony are not equally desirable for individuals from different cultures or between men and women, Colston and Lee (2004) argue that it is possible that "(a) verbal irony accomplishes a variety of pragmatic functions that are different

from those of other figurative and indirect forms, and (b) males, due to whatever causal mechanisms, might have a different set of discourse goals on average than females" (p. 291). According to this explanation, men's discourse goals match the pragmatic functions of verbal irony (such as enhancing the condemnation expressed towards others) better than this is the case for females' discourse goals. Furthermore, as the use of irony poses the relatively greater risk of misinterpretation (as compared to literal communication), Colston and Lee (2004) explain men's higher (vs. women's lower) endorsement of irony by men's higher risk taking tendencies (relative to women's lower risk taking tendencies).

Conclusion

Although their investigation did not support the "discourse goal match" explanation, Colston and Lee's (2004) study can be argued to illustrate how the present thesis can advance the field of irony research. Their aim was to explore why particular social and cultural differences are found in nonliteral language use. By looking at gender differences, they used a *known groups paradigm*. That is, in order to gain a deeper understanding of why irony is used (i.e., the pragmatic functions of irony), they compared two groups known for different discourse goals (i.e., men and women). Of course, this approach can be seen as limited, as men and women differ in many more respects than the variables assumed to be associated with gender differences in irony use (i.e., discourse goals and risk taking). However, Colston and Lee's (2004) approach can be taken as a model for the present thesis when it comes to learning more about the phenomenon of verbal irony by finding out which personality traits and abilities are associated with irony detection and use. In the same way in which the gender differences approach was construed as a way to learn more about the discourse goals of irony in Colston and Lee's (2004) study, it may be argued that the individual differences approach used in the present thesis may help to gain knowledge about functions of irony. To illustrate: Averbeck and Hample (2008) attempted to explore "why we produce ironic messages" by

linking irony use to aggressiveness. As they argued, the finding that irony use correlates with aggressiveness indicates that irony could be a means of interpersonal aggression (cf. Averbeck & Hample, 2008). It is noteworthy that the rationale of their investigation can be seen as different from the logic mostly found in studies in personality research: commonly, it is the personality constructs that are supposed to be validated by comparing known (e.g., extreme) groups with natural differences as to certain phenomena, and not the other way around; i.e., usually it is not the assumed phenomena (e.g., the aggressive function of irony) that are supposed to be validated by use of a personality construct (e.g., aggressiveness). Of course, there may be confounding variables between aggressiveness and irony use. But it can be argued that the outlined individual differences approach to answering the question of whether aggression is a function of irony provides still higher internal validity than answering the same question by, for example, comparing men and women (i.e., under the presupposition that there are gender differences in aggression between men and women). Irony as a communication phenomenon is assumed to serve various intra- and interpersonal functions (see section "Why do we use irony?"). Accordingly, as different personality traits can be argued to go along with different experiential and behavioral tendencies in the social domain, learning more about which personality traits are associated with irony detection and use may help to learn more about which experiential and behavioral tendencies are relevant for irony. To illustrate: If trait bad mood was found to be associated with irony use in a positive direction (as it was hypothesized by Ruch & Hofmann, 2012), one could hypothesize that irony is used, for example, in order to express sullen, grumpy, or grouchy feelings.

Some thoughts on phylogenetic aspects of irony

As several of the theories of irony outlined earlier in the present thesis, such as the *allusional pretense theory of discourse irony* by Kumon-Nakamura et al. (1995), the *echoic reminder*

theory of verbal irony by Kreuz and Glucksberg (1989), the *pretense theory of irony* by Clark and Gerrig (1984), or the *"echoic mention" theory* by Sperber and Wilson (1981) suggest, irony can involve an allusion to an antecedent event, an overt pretense, or an imitation of something that was previously stated. By uttering something that was previously said, the use of irony may resemble the *aping* of the speaker of the antecedent utterance. For example, Person A may predict: "The dinner party will be totally boring!". When the facts prove Person A's prediction wrong, Person B may mock Person A for his or her misanthropic or pessimistic outlook by ironically uttering "Well, the dinner party was indeed totally boring!". This example illustrates that irony may involve an imitation (i.e., an echoic mention) as well as a pretense when Person B acts *as if* he or she was Person A in order to reject his or her utterance as overly pessimistic or simply not true. Like in this example, it can be argued that in some cases of irony, the ironic utterance may resemble an *aping* of someone or something that was previously uttered. According to the *Merriam-Webster Dictionary, to ape* means to mimic, copy, or imitate something or someone, and is defined to be a synonym of the verb *to mock*. As a noun, *ape* is the term for monkey, or, more precisely, for primates such as the chimpanzee, gorilla, orangutan, or gibbon. In the German language, the closest translation of the transitive verb *to ape* is *nachäffen*, which again is rooted in the term for the animal species of the ape (i.e., *Affe*). As the meaning of *nachäffen* in German that goes beyond the meaning of *to ape* is best translated using the English terms *to imitate* or *to mock*, the association between mocking imitation and the ape as the phylogenetically nearest neighboring species of the *homo sapiens* is also evident in the German language (and there may be even more languages in which this is the case). This observation may lead us to the question of whether *aping* as a very basic form of mocking something or someone may be a prototype of verbal irony in the repertoire of social interaction patterns and stereotypical figures of speech that have evolved in the course of the phylogenetic development of the

human species and its cultures. Although there is an ongoing debate about whether apes are in fact able to act as if they were someone else in terms of a pretense play (e.g., Gomez, 2008; Whiten, 1992), in human languages, a mocking exposure of someone's weaknesses or transgressions has been associated with an ape-like imitation of the very behavior one is trying to make fun of or criticize. It is not intended to postulate here that there is irony among apes. Rather, it can be argued that ape-like mocking imitation of someone else's behavior may be prototypical element of what has evolved as verbal irony in humans, which in a very basic form involves echoing an antecedent utterance in order to derisively "reject it as ludicrously false, inappropriate, or irrelevant" (Sperber & Wilson, 1981, p. 308). To conclude, the fact that mocking imitation in language use has been associated with primates' behavior may be seen as a hint to the possibility that *aping* is a behavior that developed early in phylogenesis. Possibly, irony as an elaborate form of derisive simulation roots in aping as a more "primitive" form of mocking imitation.

This consideration can be seen as important for the present thesis because it may support the aim to conceptualize irony detection and irony use as an ability and a behavioral tendency, respectively: If irony behaviors truly were to be understood as systematic und enduring tendencies (in terms of individual differences), it would be thinkable (if not a theoretical presupposition) that these tendencies have a genetic foundation. Hence, if there was a phylogenetic development of irony behaviors, this could be taken as an indicator for a possible heritability of irony detection aptitude and irony use tendency (or the psycho-physiological components possibly underlying these behaviors).

Assuming an individual differences perspective in irony research

Targeting the investigation of interindividual differences in irony detection and irony use implies that one assumes there to be a substantial amount of systematic interindividual

variance in these two domains of irony behaviors. In the present thesis, it is expected that this variance can be explained (to a certain degree) by personality and ability, as irony behaviors can be seen as an "arena" in which traits and abilities manifest themselves in behavior. To provide an overview of the relevant literature, this section reviews the studies that could be found to investigate irony detection or use in the population of normally functioning adults by relating them to known traits and abilities.

Postformal thinking and creativity in irony detection

As an exception to the general neglect in research of healthy adults' irony detection in relation to individual differences variables, a study by Blouin and McKelvie (2012) jointly investigated the reception and also the appreciation of irony with concepts from the realm of individual differences, namely creativity and *postformal thinking*. As Blouin and McKelvie (2012) report in their review of the relevant literature, postformal thinking transcends adolescents' formal thinking and is defined to comprise two stages which are seen as bound to post-adolescent developmental maturity. Whereas formal thinking delineates the preference for dual thought patterns and a fixed view of the world with no tolerance for contradiction (which is seen as most prevalent in young adults), the first stage of *postformal* thinking, *relativistic thinking* (most prevalent in older adolescents or young adults), is characterized by a mode of thought in which contradiction is seen an inherent aspect of circumstances. In the second stage of postformal thinking, *dialectical thinking*, individuals (typically middle-aged or older adults) reconcile contradicting aspects that arise when when making sense of the world by understanding them as facets of an underlying whole or overarching unity. As a culmination of cognitive development in dialectical thinking, the dialectical whole is constantly challenged by arising incongruities to evolve to new stages of resolution. Although Blouin and McKelvie (2012) do not explicitly declare postformal

thinking as an individual differences variable, they treat it as one when they assess it as "the ability to accept and integrate a variety to [sic!] truths that are context-dependent and may even be incompatible" (p. 41) using a preexisting measure that was previously reported to have high internal consistency and test-retest correlations. Furthermore, postformal thinking was implied to vary across adults as Blouin and McKelvie (2012) used a sample in which they observed that "the vast majority were between 18 and 23 old" (p. 47; age and gender were not recorded in their study).

In Blouin and McKelvie's (2012) study, irony detection was assessed via a task that measures the ability to identify irony and metaphor. Participants had to identify and mark irony with a pen in a dialogue involving irony (i.e., "Wow, I really thought we were beyond the 'small talk' stage, but I guess we're going to have to go back to boring old shallow conversations…", when talking about death) and metaphor (i.e., "No but seriously, Em, aren't we all ostriches with our heads in the sand?"). Additionally, they had to explain why they thought the respective part of the dialogue was irony or metaphor and, among other ratings, indicate how much they liked the respective instance of irony or metaphor. Participants' scores in the irony and metaphor identification task (which included correct identification and correct explanation of irony) correlated substantially with their scores in the task for the assessment of postformal thinking, but also with their scores in a divergent thinking task. Divergent thinking performance was assessed as a measure for creativity because the authors linked postformal thinking not only to the identification and appreciation of irony and metaphor, but also to creativity. Hence, the finding that irony and metaphor are better identified and explained by creative individuals is a secondary result in this study, and not discussed by the authors. However, as the identification of irony and metaphor were jointly scored, it is hard to distinguish whether it is the identification and explanation of irony or metaphor that is responsible for the associations found between the irony and metaphor

identification task on the one hand and postformal thinking and creativity on the other. Furthermore, it will not be discussed here whether postformal thinking can be classified as a personality trait in terms of a preference for a certain worldview, or rather as an ability to engage in holistic thinking and reasoning. Neither will it be discussed whether postformal thinking is a unique concept or a blend of different known traits or abilities. Rather, Blouin and McKelvie's (2012) findings will be taken as a demonstration of systematic variance in irony detection. As this variance is likely not to be fully explained by postformal thinking and creativity, further individual differences variables may play a role when studying the role of personality and ability in irony detection.

Rigor of decoding the implicit and *need for consistency* in irony use

As a further exception to the general neglect of an individual differences perspective in irony research, Groeben et al. (1985) investigated the role of theoretically relevant (albeit narrow) trait variables in irony production. However, presumably due to the fact that their set of studies was published in German language, their approach and their findings were not acknowledged in the English-speaking scientific community. Although Groeben et al. (1985) included irony reception in their large-scale investigation, they did not investigate the role of trait variables in irony reception. Furthermore, they did not operationalize irony reception in terms of an irony detection criterion. That means they did not strictly assess whether subjects detected the irony in the stimulus or not. Instead, they collected ratings about speaker and speech characteristics (e.g., superiority, indirectness). As regards irony use, Groeben et al. (1985) found that a set of personality characteristics were associated with irony production, namely *need for consistency* (*Konsistenzbedürfnis*) and *rigor of decoding the implicit* (*Stringenz der Implizitätsdekodierung*). Need for consistency was assessed with a preexisting questionnaire for the measurement of *tolerance of ambiguity* by Rydell and Rosen (1966;

sample items are: “A problem has little attraction for me if I don't think it has a solution.”, and “There's a right way and a wrong way to do almost everything.”). Rigor of decoding the implicit was assessed with a measure that was developed ad-hoc, consisting of riddle-stories that participants had to respond to by indicating how they made sense of the stories with open-ended responses. These traits can be seen as related to postformal thinking because *need for consistency* can be reasoned to be lower in individuals that score high in postformal thinking than in formal thinkers. Hence, the same dispositions may affect both the detection *and* the use of irony. In Groeben et al.’s (1985) study, irony use was assessed by asking participants to indicate their choice between five predefined utterances, among which one was designed as ironic. These findings are encouraging as they are a first demonstration of the role of dispositional variables in irony use and lay the groundwork for building hypotheses on the role of broad personality dimensions and humor-related traits in irony use.

The role of ability in irony detection and irony use

One of the aims of the present thesis is to jointly investigate personality and general mental ability on the one hand and irony detection on the other, in order to try to replicate and extend pre-existing studies that suggest the existence of systematic and meaningful variance in irony detection that can be explained by individual differences in terms of personality and ability variables (i.e., variance that can be explained rather than being negligible “noise”). Although it may appear worthwhile to also look at the role of intelligence in irony *use* (as irony can be seen as a witty and clever figure of speech), the role of ability in irony use will not be addressed in the present thesis for several reasons: (a) as most of the individuals who know how to use irony probably do not use irony to the full extent of their possibilities (i.e., much less frequent than thinkable), irony use can be seen as a typical behavior in of the sense of habitual preferences, (b) besides the circumstance that irony detection was found to be

correlated with intelligence, and one may hypothesize that irony use and irony detection are interdependent (as the knowledge of the concept of irony and the ability to detect it successfully may be a prerequisite for successfully using irony), there is hardly any theoretical indication for the hypothesis that general mental ability plays a role in irony use, and (c) in the conceivable case that intelligence is found to be positively correlated with irony use in a potential study, it would be hard to distinguish (albeit not methodologically impossible examine) *why* ironic utterances are preferred (vs. rejected) by more (vs. less) intelligent individuals (for example, due to enhanced irony detection in the stimuli used or due to enhanced wit), so the informative value of jointly investigating irony use and intelligence may be limited.

In contrast, it may be worthwhile to study the role of ability in irony detection for several reasons: (a) as it is expected that not all individuals can perform equally well in irony detection, irony detection performance can be seen as theoretically related to individual differences variables delineating maximal behavior, (b) hence, jointly investigating irony detection and general mental ability can help to test the assumption that there is systematic interindividual variance in irony detection (rather than merely negligible "noise") that can be explained by known individual differences variables, and (c) there is empirical support for the assumption that intelligence explains variance in irony detection. For example in Mitchley et al.'s (1998) study, intelligence was numerically more strongly associated with the errors made in ironic items (i.e., r = -.53) than with the errors made in the literal items of the irony detection measure (i.e., r = -.42) among schizophrenic patients. This difference was even more pronounced when general mental ability was assessed with a measure of premorbid intellectual ability, albeit the correlations did not reach the level of significance. Supporting this finding, Varga et al. (2014) found that irony detection was positively correlated with intelligence (i.e., r = .54) in their schizophrenic patient group. However, in

the two named studies, the association between irony detection performance and intelligence in the healthy control groups was not reported.

Therefore, intelligence will be examined as a potential correlate of irony detection aptitude in PART II of the present thesis. It is expected that irony detection and intelligence correlate with each other highly enough to confirm pre-existing findings that irony detection, as assessed with a newly developed measure, poses a cognitive demand for the individual. At the same time, irony detection and intelligence are expected to be inter-correlated low enough to validate that the construct of irony detection aptitude is distinct from general mental ability. Furthermore, ironic praise, as the more complex and less stereotypical form of irony, is expected to be more dependent on intelligence than ironic criticism (see PART II of the present thesis for details).

Assessment of irony detection and irony use

The studies conducted for the present thesis made use of both preexisting and newly developed measures for the assessment of irony detection and irony use. To enable the reader to integrate the measures that were chosen for the studies conducted for the present thesis, some of the most frequently used measurement approaches for the assessment of irony detection and irony use in the existing literature are characterized in the following sections. In addition, the shortcomings of some of the existing measures will be discussed.

Measures of irony detection

As a basic common feature of previously used measures of irony detection, participants were typically provided with a scenario ending with a target utterance. While there are studies using audiotaped stimuli (e.g., Ackerman, 1983; Dews et al., 1995; Winner et al., 1998), written stimuli are used more commonly in the existing literature. In face of the plethora of

irony detection measures that were administered in written format, it seems justified to say that irony can be detected with no other cues provided than those possible to write down. This does not mean that irony is not easier to detect when recipients can make use of paraverbal cues, such as facial expression or intonation. In fact, McDonald et al. (2006) used audiovisual stimuli that were completely deprived of factual cues, so their test-takers had to rely only on paraverbal cues, such as intonation and facial expression of the actors, when appraising videotaped dialogues. However, as the next section aims to describe, these cues are not necessarily important for irony detection. It could even be argued that audiotaped or audiovisual stimuli add noise to irony detection studies (and, arguably, especially so when comparing ironic stimuli to non-ironic stimuli), as (a) the phonological features of stimuli may be hard to control in terms of comparability, and (b) audiotaped stimuli may add confounding features to the ironic utterances.

Does irony detection require an ironic tone of voice?

It was suggested that ironic utterances are typically accompanied by certain phonological markers; however, there is no consensus as to which markers these are, and there are even contradictory assumptions about some of them (such as the pitch of voice, which some deem to be higher and some deem to be lower in ironic utterances; cf. Attardo, Eisterhold, Hay, & Poggi, 2003). Clark and Gerrig (1984) account for this phenomenon when they claim that, in support of their pretense theory, the ironic speaker is supposed to assume a voice that is appropriate for the person he or she is pretending to be while exaggerating and caricaturing this voice (see section “The pretense theory of irony by Clark and Gerrig [1984]”).

As Kreuz and Roberts (1989) suggest, the ironic tone of voice may be not be specific for irony, but rather specific for *hyperbole* (i.e., a choice of words exaggerating the true circumstances). As they suggest, hyperbolic statements are typically expressed by heavy

stress and a slow speaking rate, which has been defined as the ironic tone of voice in previous literature (cf. Kreuz & Roberts, 1989). In other words, the ironic tone of voice may be confounded by the use of hyperbole in ironic utterances. In a similar vein, Sperber and Wilson (1981) mention the existence of an ironic tone of voice. However, they suggest that the ironic tone of voice is not specific for irony, but rather for the emotional valence of the attitude expressed by the ironic utterance when they point out that the ironic tone of voice "is merely one of the variety of tones (doubtful, approving, contemptuous, and so on) that the speaker may use to indicate his attitude to the utterance or opinion mentioned" (Sperber & Wilson, 1981, p. 311; see also section "The 'echoic mention' theory by Sperber and Wilson [1981]" of the present thesis). In line with this conclusion, Bryant and Fox Tree (2005) did not find support for the postulation that an ironic tone of voice, which they defined as *prolonged articulation and exaggerated pitch*, has an impact on the degree of sarcasm perceived in ironic utterances.

Characterizing existing instruments for the assessment of irony detection

Typically, in the existing literature, the stimuli used for the assessment of irony detection consist of a list of short *stories* (oftentimes also called *scenarios*, sometimes also labeled *vignettes*) providing information about a situation with two or more protagonists interacting. The scenarios usually end with a target utterance that is designed as either ironic or non-ironic. Ironic target utterances can commonly be characterized as counterfactual statements that the speaker in the scenario intends to be interpreted as such by the addressed protagonist. For the irony to be detected, the reader of the scenarios has to realize that the utterance is counterfactual to the information provided previously in the scenario. Likewise, the reader has to realize that the speaker in the story knows that the recipient is aware of the true circumstances and accordingly he or she (the recipient) has to appraise the utterance as

counterfactual. Importantly, the reader also has to be aware of the fact that the recipient knows about the speaker's knowledge of the recipient's knowledge. In other words: in order to detect the irony, the reader has to appraise that there is a common ground of knowledge about relevant facts between the conversation partners that is openly violated by the ironic utterance. Although the target utterance mostly represents a mock positive evaluation of negative circumstances described in the scenario (i.e., *ironic criticism*; cf. Sperber & Wilson, 1981), there are also studies using mock negative evaluations of positive circumstances (i.e., *ironic praise*; e.g., Langdon, Davies, & Coltheart, 2002). Characteristically, the ironic utterance refers to a positive expectation (i.e., in the case of ironic criticism) or negative expectation (i.e., in the case of ironic praise) that was not fulfilled (for the critical nature of irony see, for example, Garmendia, 2014; Sperber & Wilson, 1981). A special case can be seen in Winner et al.'s (1998) stimuli (that were used for Study 3 of Part II of the present thesis), as they defined irony as a counterfactual statement that is neither ironic criticism nor ironic praise. Rather, the counterfactual (i.e., ironic) statements are uttered jokingly in order to cover up the speaking protagonists' embarrassment over getting caught during a violation of some social rule. As the ironic speakers are aware of the fact that their conversation partners know about their "sneaky" transgression, they overtly pretend that they did not do it with the aim of producing a certain humorous and embarrassment-reducing effect. This is why it was possible to rephrase "using irony" with "joking" when asking participants about their interpretation of the stimuli in Winner et al.'s (1998) study.

As regards the response format of the instruments used for the assessment of irony detection, there is a fair amount of different approaches to assess whether participants choose a literal or an ironic interpretation of the stimuli. An overview of different measures for the assessment of irony detection aptitude is given in Table 1. As Table 1 shows, many of the previous studies used fact questions in order to determine whether irony was detected or not.

Irony detection measures are typically scored by counting the number of ironic items (i.e., scenarios with an ironic target utterance) that participants' responses indicated as having been interpreted as ironic. The total number of correctly answered items or interpretation errors is then used to compute a total score (e.g., an error rate). In the remainder of the current section, different procedures for the operationalization of irony detection as used in the existing literature will be described.

One example of a measure of irony detection using fact questions is provided by Ackerman (1983). He asked his participants to answer story-specific fact questions and speaker attitude questions that revealed whether or not they appraised the utterance as literal or ironic in the given stories. Happé (1993) used a somewhat different approach to determine whether her subjects chose a literal or ironic interpretation of the stimuli. She asked her participants how the target utterances in her "metaphor vs. irony task" was intended to be understood by the speaker by providing a choice of two options, one of which was reflecting the ironic and one of which was reflecting the literal meaning of the target utterance (e.g., "What does David's father mean? Does he mean David is clever or silly?"; Happé, 1993, p. 119).

In order to assess whether their participants detected the irony in the stimuli, Langdon and Coltheart (2004) simply asked them whether it made any sense for the fictional speaker to make the target utterance. The rationale of this procedure stems from the circumstance that, due to their counterfactual nature, ironic utterances semantically do not fit in with the situational context, such as the contextual information provided in the respective stimulus stories, if one does not grasp the ironically embedded meaning.

Table 1

Overview of Measures for the Assessment of Irony Detection and Irony Use.

	Participants' task	*Stimuli*
Irony detection measures		
Winner et al. (1998)	Categorize a target sentence as "irony" or "joking".	Short stories with target utterances
Ackerman (1983)	Answer dichotomous (i.e., "yes" or "no") story-specific fact questions and speaker attitude questions.	
McDonald and Pearce (1996)	Explain what they thought was going on between the speakers (responses are coded dichotomously along a list of criteria).	
Happé (1993)	Answer dichotomous questions about intended meaning.	
Langdon et al. (2002)	Answer dichotomous questions as to whether it made sense for the speaker to make the target utterance.	
Channon, Pellijeff, and Rule (2005)	Explain verbally the final remark, the action or the event immediately after presentation of the scenario (responses were scored by raters).	
Irony use measures		
Averbeck and Hample (2008)	*Rating-based*: indicate on a rating scale the likelihood with which they would make each one of ten predefined ironic remarks.	Scenarios and utterances
Matthews et al. (2006)	*Forced choice*: indicate which of two predefined utterances they would choose in response to the given scenario.	
Dress et al. (2008)	*Free production*: generate a free response that they would make if they were in the place of one of the described protagonists in the respective short story.	Short stories
Hancock (2004)	*Face-to-face interaction*: discuss celebrity fashion images with their conversation partners (that were designated examples of poor fashion taste) as if they were commenting a fashion show (e.g., by praising the outfits). Conversations were scored by raters using a coding scheme.	Discussion topics
Ivanko, Pexman, and Olineck (2004)	*Self-report*: answer questions assessing their estimate of the likelihood with which they use sarcasm.	Self-report questions

Irony detection has also been operationalized by assessing participants' judgments of whether a non-ironic paraphrase of the target utterance in the scenario is true or false. For example, Ackerman (1982) used two kinds of questions that had to be answered in a *true or false* format. One kind of question asked about the speaker's state of mind or propositional attitude that the speaker intended to communicate (e.g., displeased, angry, or hostile). The other kind of questions asked if the non-ironic "translation", in terms of a literal paraphrase of the target utterance, was true or not. In a similar vein, McDonald and Pearce (1996) made use of questions targeting mental states and emotions of both the speaker and the target of the ironic utterance.

Other studies used open response formats. For example, participants had to verbally respond to the question of what the fictional character in the stimuli meant by his or her remark in a face-to-face assessment with the experimenter, who scored participants' responses as to whether they reflected literal or ironic interpretation of the target stimulus (e.g., Channon et al., 2005).

Some studies assessed irony detection by rewording the use of irony as "joking" and asking participants whether the ironic utterance was meant as a joke (i.e., irony) or a lie (e.g., Winner et al., 1998). The instrument introduced by Winner et al. (1998) was also used by Gaudreau et al. (2013). As they explain, like Winner et al. (1998) they avoided the term "irony" in their materials as not to prime their participants. As Gaudreau et al. (2013) did not elaborate on this issue, the next section will deal with the question of what forms of priming can occur during the assessment of irony detection. .

Indirect response formats in the assessment of irony detection

This leads to the report of another observation about the existing measurement methods for the assessment of irony detection: Usually, participants are not asked immediately whether

they think that the stimuli are ironic or not. Rather, irony detection is assessed indirectly or *unobtrusively*. The reason for not making irony salient and thereby making participants aware of the possible occurrence of irony can only be guessed at, as usually they are not reported in the respective studies. As already mentioned, the notion that it is advisable to avoid priming effects by unobtrusive assessment of irony detection can be found in the literature (Gaudreau et al., 2013). Avoiding "priming" by distracting from the true intention of an irony performance task may have at least two benefits: (a) given that the chance hit rate in most of the instruments can be expected to be at 50%, since there are only two options, each specifying the intentions of the speaker, making irony salient may lead to ceiling effects when tasks become too easy to solve, and (b) especially when comparing cognitively impaired groups, clinical participants may produce socially desired responses in terms of demand effects, as they may guess that irony detection may be affected by their disorder (especially when no non-ironic control stimuli are used). Since it can be seen as an important aspect of the assessment of irony detection, the role of irony alertness will be revisited in PART II of the present thesis. Based on the findings of Study 3 in PART II, it will be suggested that, for the investigation of the role of personality and ability in irony detection, making irony salient for test-takers (i.e., an *irony alert* mode of testing) may be preferred to a procedure in which irony is not mentioned to participants (i.e., an *irony non-alert* mode of testing)—at least when using psychometrically difficult (i.e., ambiguous) stimuli.

A new direction in the assessment of irony detection

As the present thesis innovates the field of irony research by also addressing healthy adults' irony detection (rather than resorting to clinical groups), new means of assessment that are tailored to the performance level of the target group are needed. That is, an ideal test of irony detection performance would allow for discrimination between different levels of high irony

performance. A novel test is required because pre-existing measures typically suffer from ceiling effects (i.e., due to low psychometric difficulty of the items). Such ceiling effects can be found, for example, in normally functioning control groups in studies comparing them to clinical groups (indicated by close-to-perfect detection rates and very low standard deviations (e.g., Langdon et al., 2002; Mitchley et al., 1998). To fill this research gap, newly developed instruments tailored to normally functioning adults' performance levels will be introduced in PART II of the present thesis.

Assessment of irony use

While there are, remarkably, fewer studies on irony *use* than on irony *detection* in the existing literature, irony use has nonetheless been operationalized by using a whole variety of different approaches. An overview of different measures for the assessment of irony use tendency is given in Table 1. As Table 1 shows, generally, these approaches fall into five broad categories: (a) rating-based measures, (b) forced choice measures, (c) free production measures, (d) face-to-face interaction tasks, and (e) self-report questionnaire. As an example for a *rating-based measure*, Averbeck and Hample (2008) asked their participants to indicate (on a rating scale) the likelihood with which they would make each one of ten predefined ironic remarks in a given scenario. In one of the experimental conditions in this study, participants were instructed to assume that they were sharing a flat with a noisy, messy, parasitic, and disrespectful person. Participants had to indicate how likely it is that they would address their flatmate with a list of 20 criticisms, of which ten are phrased as ironic and ten as literal. This measure is used in one of the versions provided by Averbeck and Hample (2008) and described in more detail in PART III of the present thesis.

As a measure falling into the *forced choice measure* category, the procedure used by Matthews et al. (2006) uses short stories in which two protagonists are interacting.

Participants were requested to indicate which of two predefined utterances they would choose in response to the given scenario if they were to encounter the described situation in real life, from the perspective of the protagonist making the final utterance. One of the utterances provided in the selection task was designed as ironic and one as a literal response. Scores reflect the frequency with which participants choose an ironic statement. This measure is adapted from Matthews et al. (2006) for the use for the present thesis and described in more detail in PART III.

As regards the category of *free production measures*, a study by Dress et al. (2008) used an open-ended response format by employing a list of short stories designed to elicit an ironic utterance. Participants had to picture themselves as taking part in the interaction described in the short stories (which were adapted from the stimuli introduced by Kreuz & Glucksberg, 1989) and generate a free response that they would make if they were in the place of one of the described protagonists in the respective short story. Dress et al. (2008) hoped to elicit ironic responses by the following setup: In the scenarios, one story character makes a positive prediction that then fails to come true, with the other story character being a witness (or even being affected by the negative consequences of the failing prediction). The rationale of this procedure can be seen in the fact that the positive prediction provides material that can be mentioned in order to convey a critical attitude towards the false prediction or the failing expectation. That is, when participants were requested to respond as if they were the other story character, they could use a mock positive evaluation of negative circumstances in terms of an ironic criticism in response to these stimuli (cf. Sperber & Wilson, 1981).

As an example of the *face-to-face interaction tasks* category, Hancock (2004) used two conversation tasks. Similar to Rockwell and Theriot's (2001) procedure (see section "Culture and gender differences in irony use"), the tasks were designed to elicit irony by

making participants utter something positive about negative circumstances. For example, in one task participants had to discuss celebrity fashion images with their conversation partners that were designated examples of poor fashion taste, as they were taken from the "Worst Dressed" pages of pop-culture magazines. Participants were instructed to act as if they were commenting a fashion show (e.g., by praising the outfits). The conversations in the face-to-face condition of Hancock's (2004) study were videotaped and scored using a coding scheme incorporating four types of speech acts in which irony can occur as mentioned by Gibbs (2000): (a) sarcasm (i.e., stating the opposite of the intended meaning in order to convey a negative attitude), (b) understatement (i.e., counterfactual statement in terms of stating less than is the case), (c) hyperbole (i.e., exaggeration of the situation), and (d) rhetorical questions (i.e., when the speaker seemingly asked a question in order to convey his attitude but did not await an answer).

As regards the *self-report* category, Ivanko et al. (2004) used a 16-item questionnaire to assess how participants appraise the likelihood with which they use sarcasm *in general* (e.g., "How sarcastic do you think you are?") and the likelihood with which they would use sarcasm in certain *specific* predefined situations (e.g., "You just got a big promotion at work. You are having dinner with your family to celebrate your achievement…") on a 7-point scale (*not at all likely* to *extremely likely*).

Taken together, there are various forms of measures for the assessment of irony use. As a general weak point, the identification of irony use can be seen as being subject to the respective researchers' appraisal of whether a given response is ironic or not in all of the described approaches. That is, the respective researchers either designed predefined responses as ironic or scored the occurrence of irony according to their own appraisal of free speech productions. As a proxy for ironic intent, responses are commonly defined as ironic if the valence of the literal interpretation of the utterance is opposite to the valence of the true

circumstances. However, strictly speaking it would be necessary to ask participants explicitly whether they intended to use irony or not, at least as a control question after the irony use assessment (as not to bias participants' responses in terms of *irony alertness* during the irony use measure). As none of the described studies used these kinds of control questions, it is not clear how much *noise* should be expected in the responses of the described irony use measures, such as noise caused by guessing answers if participants cannot make sense of the scenarios used (although it would be possible to estimate the proportion of error variance in the score by testing the reliability of the measures).

Aiming to derive a theoretical model of individual differences in irony behavior

As a secondary objective of the present thesis, it is aimed to ultimately derive a theoretical model summarizing and organizing individual differences in irony behavior. This model will deal the individual differences constructs relevant to irony behavior that are tested or theoretically assumed in the present thesis. As a newly introduced construct, the *sense of irony* as an individual differences variable manifesting itself in the extent to which people detect irony, produce irony, seek irony, or enjoy irony (see PART I of the present thesis) will play a prominent role in this model. According to Cronbach and Meehl (1955), a theoretical construct is constituted by a system of *laws*, which they refer to as a *nomological network*. As Cronbach and Meehl (1995) point out, "'[l]earning more about" a theoretical construct is a matter of elaborating the nomological network in which it occurs, or of increasing the definiteness of the components." (p. 290). In the General Discussion section of the present thesis, an attempt will be made to summarize the "laws" in which (a) observable properties or quantities of the *sense of irony* relate to each other, (b) the *sense of irony* relates to observables; and (c) different theoretical constructs relate to the *sense of irony* (cf. Cronbach & Meehl, 1955). This step can be seen as crucial when assuming systematic interindividual

variance in irony behavior due to the *sense of irony* and ultimately targeting its assessment, as "[t]o validate a claim that a test measures a construct, a nomological net surrounding the concept must exist" (Cronbach & Meehl, 1955, p. 291).

Summary and conclusion of General Introduction

This section summarized the current knowledge on irony behavior in terms of pre-existing theories of irony, attempting to explain irony phenomena by making assumptions about which general features help to distinguish irony from non-irony (such as a negative attitude toward a failed expectation, e.g., Utsumi, 2000), which linguistic characteristics typically go along with irony (such as "echoic mention", e.g., Sperber & Wilson, 1981), or which pragmatic processes are involved in irony detection (e.g., the "cooperative principle", Grice, 1975). It was described how these theories relate to the aim of the present thesis, in that they indicate which traits can be expected to be relevant for the prediction of irony behaviors, and also which aspects should be considered for their assessment. It is noteworthy that none of the existing theories make assumptions about whether characteristics of the speaker play a role in irony behavior, indicating that the present thesis addresses a gap in research by aiming to introduce a theory of irony reception that incorporates an individual differences perspective. Next, as two central functions of irony, *humor* and *aggression* were found as pervasive in the literature. Since irony may be involved in both the phenomena of humor and aggression, irony may have a special role when it comes to the disparaging part of the humor spectrum. Therefore, in addition to personality dimensions relating to aggression (i.e., psychoticism) the present thesis will take on humor-related traits that not only delineate playful compassionate aspects of the sense of humor, but also those traits that go along with the preference for laughing-at humor.

Furthermore, the clinical approach in irony research was examined in the introduction. This approach utilizes patient groups suffering from a depletion of the abilities denoted by *Theory of Mind*. The description of this approach was viewed as worthwhile to include because it demonstrates the potential advantage of investigating individual differences in irony detection, an approach that can overcome the limitations of clinical studies (such as problems concerning internal validity and feasibility) by targeting the population of normally functioning adults. Next, studies investigating culture and gender differences were reported. These studies compared groups known for differing discourse goals (e.g., men and women, e.g., Colston & Lee, 2004) in order to gain a deeper understanding of why irony is used (i.e., the pragmatic functions of irony). Of course, this approach can also be seen as limited because, for example, men and women differ in many more respects than the variables assumed to be associated with gender differences in irony use (i.e., discourse goals and risk taking, cf. Colston & Lee, 2004).[3] Previous studies that reported initial findings as to the role of creativity and intelligence in irony detection and use were discussed in the General Introduction to demonstrate that it is indeed worthwhile to assume an individual differences perspective in irony research. In addition, the assessment of irony detection and use was addressed by describing the features of the most established measurement procedures. Lastly, the aim of the present thesis to derive a theoretical model of individual differences (and especially so the assumed *sense of* irony) as predictors of irony behavior was described. This

[3] However, Colston and Lee's (2004) approach of drawing conclusion from correlations between gender and irony use to learn more about the phenomenon of irony can be taken as a paradigm for drawing conclusions from the expected results of the present thesis from an interdisciplinary angle: Learning more about *who* is able to detect irony and inclined to use irony (i.e., characterizing "the typical irony user" in terms of personality and ability) may also help research disciplines other than the individual differences field (such as experimental psychology) to generate and test hypotheses about *when* and *why* irony is detected and used. Eventually, pre-existing assumptions and knowledge about the "laws" in which irony occurs could in futures works be reappraised and extended using an interdisciplinary point of view.

model is intended to be used as a basis for deriving a theory explaining successful irony reception.

If one main conclusion is to be drawn from the General Introduction, it's that the literature reviewed therein indicates that the present thesis is targeting an important gap in research. Being first to target the question whether (a) systematic interindividual differences in irony performance can be found and (b) whether this interindividual variance can be explained by personality and ability variables, the present thesis can help to get a picture of *who* is able or inclined to detect or use irony. The aim of dealing with these questions, is to open up a new field of research for both personality and irony research, by conceptualizing irony detection as an aptitude and irony use as enduring tendency for which, according to a compelling (yet neglected) pre-existing theoretical and empirical indication, a link to ability and personality should be expected. As a by-product of the results expected in PART II and PART III of the present thesis, pre-existing assumptions about the special role of humor-related traits will be put to the test by (a) exploring their role in irony detection as compared to the role of broad personality dimensions (i.e., the Big Five dimensions of personality) and (b) testing their incremental predictive value for irony use beyond broad personality dimensions (i.e., Eysenck's personality dimensions).

Previous experimental findings are viewed as encouraging for the aim of the present thesis to introduce irony behavior as a phenomenon native to the realm of individual differences. Since substantial relations to general mental ability and the affective, behavioral, cognitive, and motivational aspects of personality traits can be expected, irony behaviors should not remain a topic of experimental research (or the study of cultural groups or gender) because—as a central claim of the present thesis—(a) systematic interindividual variance in irony detection and use exists and should be described and (b) this interindividual variance

cannot be explained sufficiently while neglecting ability and personality as two central domains of individual differences.

Interindividual Differences in Verbal Irony Detection and Use: The Role of Personality and Ability

PART I. The aim of the paper is fourfold: (1) demonstrate why humor scholars should study irony, (2) explore the need for considering interindividual differences in healthy adults' irony performance, (3) stress the necessity for developing tools assessing habitual differences in irony performance, (4) indicate future directions for joint irony and humor research and outline possible applications.

PART II. The paper aims to distinguish two facets of irony detection aptitude, namely the aptitude to detect ironic criticism vs. the aptitude to detect ironic praise, in order to allow for more accurate predictions and conclusions when exploring personality and ability correlates of irony detection aptitude. As a methodological prerequisite, the aim is to develop a novel test for the assessment of the detection of ironic criticism and ironic praise. Because irony detection aptitude is assessed indirectly, and because difficult items are used in order to prevent ceiling effects, the materials will be evaluated by using an experimental approach, and also by including additional explicit ratings of ironic content.

PART III. The aim of this paper is to explore whether (a) broad personality dimensions and humor-related traits predict irony use and, if this is the case, whether (b) humor-related traits have incremental explanatory value over broad personality dimensions in this prediction. It is hypothesized that traits facilitating humor-related behavior (i.e., gelotophilia, katagelasticism, the aggressive and the affiliative humor style, and the histrionic self-presentation style) as well as trait bad mood will predict irony use in a positive direction, just as traits that impede humorous behavior (i.e., trait seriousness) will predict the use of irony in a negative direction.

PART I

Virgin soil in irony research: Personality, humor, and the "sense of irony"

Richard Bruntsch, Jennifer Hofmann, and Willibald Ruch

Bruntsch, R., Hofmann, J., & Ruch, W. (2016). Virgin soil in irony research: Personality, humor, and the "sense of irony". *Translational Issues in Psychological Science, 2*(1), 25-34. http://dx.doi.org/10.1037/tps0000054.

Abstract

The aim of the paper is fourfold: 1) show why humor scholars should study irony; 2) explore the need for considering interindividual differences in healthy adults' irony performance; 3) stress the necessity for developing tools assessing habitual differences in irony performance; 4) indicate future directions for joint irony and humor research and outline possible applications. Verbal irony is often employed with a benevolent humorous intent by speakers, but can also serve as a means of disparagement humor. In both cases, encoding and decoding activities entailing irony need to be considered in the context of the psychology of humor. We argue that verbal irony performance can be considered a phenomenon native to the realm of humor and individual differences. We point out that research has widely neglected the meaningfulness of variance in irony performance within experimental groups when looking at determinants of irony detection and production. Based on theoretical considerations and previous empirical findings we show that this variance can be easily related to individual differences variables such as the sense of humor, dispositions towards laughter and ridicule (e.g., gelotophobia), and general mental ability. Furthermore, we hypothesize that there is an *enduring trait* determining irony performance we will label the *sense of irony*. The *sense of irony* possibly goes along with inclinations towards specific affective and cognitive processing patterns when dealing with verbal irony. As an application, novel irony performance tests can help to study psychological and neuro-physiological correlates of irony performance more feasibly, i.e., in non-clinical groups.

Keywords: Assessment, Gelotophobia, Humor, Individual Differences, Irony

Introduction

When we use verbal irony, we typically utter something different from what we actually want to express. For example, we utter an opposite of what we mean or use assertions that are counterfactual, oftentimes in order to communicate a critical attitude (Garmendia, 2014; Haverkate, 1990). Characteristically, we expect the listener to get the intended meaning of what we say nonetheless (Groeben & Scheele, 2003). The category of verbal irony typically entails positive evaluations of negative circumstances (e.g., ironic criticism via a mock compliment) as well as negative evaluations of positive circumstances (e.g., an ironic compliment via a mock criticism), with the latter being viewed as less prototypical for verbal irony (for an account on this asymmetry, see for example Kreuz & Link, 2002). Although a single example has to fall short of representing verbal irony in its variety, it might be illustrating to take a look at the following instance provided by Gibbs (1986, p. 8): “Gus just graduated from high school and he didn't know what to do. One day he saw an ad about the Navy. It said that the Navy was not just a job, but an adventure. So, Gus joined up. Soon he was aboard a ship doing all sorts of boring things. One day as he was peeling potatoes he said to his buddy, "This sure is an exciting life."”. Here, the reader is expected to grasp Gus’ intention to state that his life as a soldier is boring (with the possible subtext of a critical attitude toward the false promise made by the advertisement as a meta-message).

Although speakers usually want their irony to be recognized, listeners do not always detect it, i.e., they get the meaning of the ironic utterance wrong. Differences in irony detection performance were linked to characteristics of the stimuli or context, age-related developmental stages, or attributed to pathological or abnormal cognitive impairment (for an overview, see for example Colston & Gibbs, 2007). Individual differences in terms of a variation of maximal and typical irony detection (and production) performance, stable across situations and time, have been neglected. Therefore, to date, there is only scarce evidence to answer the question of who is inclined to use ironic speech and who is inclined to get the

meaning of ironic utterances wrong among healthy adults. Following the notion that schizophrenics' impaired irony comprehension may be found also in subjects with vulnerability for psychotic illness, Langdon and Coltheart (2004) linked schizotypal personality to irony detection performance in non-schizophrenic young adults. Furthermore, some of the rare studies reporting on personality constructs considering the use of irony deal with the histrionic self-presentation style (Renner, Enz, Friedel, Merzbacher, & Laux, 2008). 2008; Renner & Heydasch, 2010). Other relevant traits have been overlooked, namely the fear of being laughed at (gelotophobia; Ruch & Proyer, 2008), and the sense of humor (see Martin, 1998).

The question arises, whether there is stable variance in irony performance that cannot be associated with known traits and abilities. In this case, a trait-like *sense of irony* can be hypothesized as an antecedent of irony performance[4]. Furthermore, the "sense of irony" can be reasoned to moderate the impact of the sense of humor on irony performance. For example, without a repertoire of acts of verbal irony (i.e., irony production), humorous intent cannot express itself via verbal irony.

The aim of the current theoretical position paper is fourfold: First, we show why it is worthwhile for humor scholars to study irony. Second, we explore the need for considering interindividual differences in irony performance in the light of the literature. Third, we stress the necessity for developing assessment tools tapping into the variance in healthy adults' irony performance in terms of both, the typical level of irony activities as well as maximal irony performance in its different facets (i.e., misses and false alarms). Forth, we indicate future directions for joint irony and humor research and outline possible applications of the consequences of assuming an individual differences perspective in irony research.

[4] Accordingly, irony performance can have both, the status of a dependent or independent variable, reliant on whether it is treated as a criterion predicted by known personality traits or as an indicator of the *sense of irony*.

Why bother about verbal irony in the psychology of humor?

Humor frequently results as a (by-) product of irony (Garmendia, 2014) and irony is viewed as used when people want to be humorous (Roberts & Kreuz, 1994). Thus –albeit humor is not necessarily involved in every instance of verbal irony (and vice versa)– humor-related individual differences possibly explain differences in irony performance. In an early analysis of the English and German language, irony was shown to emerge as a pin in a framework of terms constituting the lexical field of *the comic*, along with humor in its narrow aesthetic meaning (Schmidt-Hidding, 1963). In the academic tradition of the aesthetic, humor "is simply one element of the comic –as are wit, fun, nonsense, sarcasm, ridicule, satire, or irony– and basically denotes a smiling attitude toward life and its imperfections: an understanding of the incongruities of existence" (Ruch, 1998, p. 6). It is noteworthy that *the comic* here is defined as the faculty able to make one laugh or to amuse whereas the "other major terminological system, largely endorsed by current Anglo-American research (and in everyday language) uses humor as the umbrella-term for all phenomena of this field" (Ruch, 1998, p. 6). In other words, the term humor replaced *the comic* in language use. There are two conclusions to draw from this: first, in the terminological system of the aesthetic, irony and humor belong to the same faculty –*the comic* (as distinguished from other aesthetic qualities, such as the tragic); second, if the term humor delineates what used to be subsumed under the term of *the comic*, irony can be viewed as a humor phenomenon able to make us laugh or to amuse. Hence, irony performance may be both constituting humor as well as depend on a sense of humor. Furthermore, we argue that irony is structurally similar to humor, as both entail ability and preference components. In support of these ideas, studies aiming at assessing discourse goals or pragmatic functions of ironic communication characteristically provide humor as a rating category. For example, Dews, Kaplan, and Winner (1995) defined humor along with status elevation, aggression, and emotional control as a social function of irony and found ironic criticisms and compliments to be rated as funnier than their literal

counterparts. In a similar vein, Gibbs (2000) found jocularity ("where speakers teased one another in humorous ways", p. 12) to be more frequent (50%) than sarcasm (28%; "where speakers spoke positively to convey a more negative intent", p. 12) in five types of irony deducted from a corpus of recorded ironic conversational turns. In line with these scholars' ideas and findings, also laypeople indicated in interviews that humor is a component of verbal irony: when Roberts and Kreuz (1994) asked their participants to indicate the reasons why an individual might use irony, 65% generated a response falling into the category "to be humorous".

From the *listener's* perspective, verbal irony is often involved in speech acts experienced as humorous: when Akimoto et al.'s (2014) subjects indicated the degree of experienced humor in different kinds of target statements, utterances were rated as more humorous when they were ironic rather than literal. Furthermore, ironic turns in conversations were viewed predominantly as humorous and regularly responded to with laughter (Gibbs, 2000). Thus, humor is frequently targeted in studies on irony and non-literal speech.

In psychological humor research, on the other hand, verbal irony is still neglected. Although Ruch (1998) suggested that for research purposes humor can be used "as an umbrella term […] including negative forms of humor, since the term now tends to exclude less benevolent forms of the comic like sarcasm, mock, ridicule, satire, irony" (p. 11), irony is seldom included in humor studies. Yet, as Garmendia (2014) reasons, ironic utterances follow the structure or processes described in humor theories. Ironic utterances can be argued to entail elements of incongruity and resolution (see Suls, 1972, for accounts on incongruity), or also superiority (see Ferguson & Ford, 2008, for an overview). It was also discussed whether the humorous quality of irony roots in the "incongruity between what speakers semantically state and what they ironically imply" (Gibbs, Bryant, & Colston, 2014, p. 585). This links irony to elements of incongruity-resolution theories in humor. Supporting the assumption of superiority mechanisms, Schmidt-Hidding (1963) characterized irony by the speaker's intent

to create a mutual sense of superiority towards a third with an initiated audience and a behavior towards the next that is described as “mocking the stupid” (p. 51). In the face of these considerations, it is thus highly surprising that irony is largely absent from accounts on humor production or as a style to be humorous (i.e., “I make ironic criticisms to tease my friends”). In similar vein, irony is rarely represented in tools assessing the sense of humor[5].

Moreover, verbal irony can be reasoned to be one tool in disparagement humor (Zillmann, 1983), aiming at ridiculing, humiliating, or putting others down (i.e., target its objects with biting criticism and make others laugh at the same time). As an empirical support for this view, Leggitt and Gibbs (2000) found their participants to rate a speaker feeling more „scornful, disdainful, and contemptuous“ (p. 6) when making statements classified as sarcastic and ironic, compared to other types of statements. Colston (1997) presents findings demonstrating that ironic criticism enhances condemnation compared to literal criticism, further indicating that irony can serve as a means of “salting a wound” and making it a form of speech suitable to ridicule others. Given that ridicule typically also employs disparaging humor to critically point to or even punish social transgressions (e.g., Bergson, 1924; Titze, 2009), the “clash” of a critical attitude and humor in irony (Garmendia, 2014) can be reconciled by separating the addressees of the critical and the receivers of the humorous component. Supporting this notion, Gibbs (2000, p. 10) states: “In some cases, then, ironic comments can be both humorous and negative, precisely because people find amusement in disliked targets being disparaged”. Furthermore, if listeners do not get the irony in a conversation, this can lead to taunting amusement in the person making the ironic statement as well as bystanders.

[5] For an exception, see items of the State-Trait-Cheerfulness Inventory (trait form, STCI-T<106>; Ruch et al., 1996, item 86: “Irony doesn’t suit me”) as well as the Humorous Behavior Q-sort Deck (HBQD; Craik, Lampert, & Nelson, 1996, item 57: “Is sarcastic.”). Furthermore, the As-If-Scale (Renner et al., 2008) accounts for both humor and irony when paraphrasing use and employment of cues for verbal irony (e.g., item 2: “When I say something I often change my voice to indicate that I do not really mean what I say“).

To conclude, irony can be a *tool* to produce humorous remarks, which can be of benevolent or disparaging nature. In other words, it may depend on a person's sense of humor how inclined he or she is to use irony or to get the meaning wrong in humorous ironic utterances. Consequently, not all of us may be able or prone to use irony to the same extent, just as not everyone may have the same inclination to detect verbal irony. However, little is known about how typical and maximal verbal irony behavior varies across individuals. Therefore, we next explore the conceptualization of irony from an individual differences perspective.

Is there individual differences variance in irony detection and production?

Verbal irony has been studied in a variety of experimental settings as regards the *comprehension* (see Colston & Gibbs, 2007, for an overview) and to a lesser extent also the *production* (Averbeck & Hample, 2008; Hancock, 2004; Rockwell & Theriot, 2001) of ironic utterances. The typical experimental approach consists of systematically varying target sentences in vignettes (Kreuz, 2000). As Kreuz (2000) points out, this has been used to test narrow predictions made by competing theories on irony processing. Still, "… variables such as personality and culture remain largely unaddressed" (Kreuz, 2000, p. 105). In an outlook on the future of irony studies Gibbs and Colston (2007) state that it is not yet clear "as to what personality characteristics make someone more prone to speaking ironically" (p. 590). The authors point to the finding that participants used speaker occupation as a cue when assessing speaker's tendency to use ironic speech (Pexman & Katz, 2001, cited in Gibbs & Colston, 2007). There is even more literature dealing with the role of speaker characteristics in listeners' interpretation of verbal irony (for an overview, see Pexman, 2005). We agree that these findings –that Pexman (2005) explains using theories of irony– are a starting point for building hypotheses on who is more inclined to use irony; however they were not a strict or direct test of who uses irony. Rather, in the studies reviewed by Pexman (2005) speaker characteristics usually were either treated as an experimentally varied cue, or participants

were asked to rate how likely they think speakers of different occupations are to use ironic speech. Supporting the view that not all of us are equally likely to use verbal irony, Gibbs and Colston (2007) suggest conducting case study analyses of individuals who are renowned for their use of irony in order to explore the qualities of what they call "the ironic mind" (p. 590). Following this notion, we assume that irony performance has a stable variance across individuals with some more prone to use and understand verbal irony and some less. In other words: "the ironic mind" can be hypothesized to be an individual differences variable to be found in all of us to a varying degree.

To our knowledge, only two studies have investigated interindividual variance in irony production behavior in a targeted fashion (Ivanko, Pexman, & Olineck, 2004; Renner et al., 2008). Ivanko et al. (2004) employed a scale for self-reported use of sarcasm[6] to account for interindividual variance in an irony production task (choice of either ironic or literal responses) and also an irony interpretation task (rating the speaker's intent). Interestingly, self-reported use of sarcasm not only explained ironic statement choice in the production task but also the ratings of speaker's attributes in the interpretation task. Although the study by Ivanko et al. (2004) is encouraging, the informative value of the findings needs to be seen as limited, most of all because irony production tendencies were assessed via self-reports and also because the interpretation task was evaluated in terms of ratings of speaker's attributes (rather than scored with an irony detection performance criterion).

[6] The terms "irony" and "sarcasm" are often used interchangeably in the existing literature. Here, we say "sarcasm" rather than "irony" if the authors used this term. However, we generally use the term "irony" as it was defined as a superordinate category entailing also sarcasm. Following Schmidt-Hidding (1963), irony is characterized by saying something differently than what is meant (but with an in-group-serving intent), whereas sarcasm indeed employs the figure of irony but in the context of hostile behavior. In line with this distinction, Gibbs (1986) refers to *The Oxford English Dictionary* when he defines that "sarcasm depends for its effect on 'bitter, caustic, and other ironic language that is usually directed against an individual'" (p. 3). However, he specifies that "it is possible to make sarcastic remarks without being ironic" (Gibbs, 1986, p. 3). Hence, not least in the context of humor, "irony" rather than "sarcasm" appears to be the term covering the broad spectrum of ironic language.

When assuming an individual differences perspective, we look at persons' tendencies across a wide range of different situations involving verbal irony. We are interested in verbal irony behavior in terms of inclinations across measurement points in time and across situations. Therefore, for our approach, specifics of the situation (e.g., context or social factors) will be considered as "noise" because they add variance to the expression of underlying traits in behavior that is not specific to the person and hence distort the true influence of a person's behavioral tendency (just like the variance accounted for by person characteristics is commonly treated as error variance in experimental studies on irony testing predictions by varying characteristics of the situation). In the remaining part of the current section, existing findings on irony performance will be explored with respect to two aspects. First, they will be evaluated in terms of the reported variance between individuals within the same experimental groups looking at standard deviations (*SD*s)[7]. Second, the attention will be turned to explained variance in irony performance. The rationale behind considering this aspect is that meaningful variance in irony performance is a necessity for explaining it as a dependent variable by independent person-specific variables.

Akimoto et al. (2014) asked their participants' to decide whether the intention of a speaker in 80 scenarios was ironic or not. The chance accuracy rate alone can be expected to be 50%. They report an average accuracy rate of 96.2% with 5.3% *SD* in their irony detection task. This is a much higher hit-rate than the accuracy found in a preliminary survey the authors conducted (80%). So in face of the fact that tasks were rather easy to solve (indicated by the high means, i.e., subjects were detecting the irony in most of the tasks) there were still noteworthy differences between subjects' performances. These findings are consistent with a further study employing a dichotomous interpretation task: Winner, Brownell, Happé, Blum,

[7] Low *SD*s (in relation to the possible range of the variable) indicate that individuals do not disperse much in terms of the deviation of their performance from the sample mean. High *SD*s indicate a broad distribution of individuals' performance scores along the possible range of the variable.

and Pincus (1998) found an average error rate of .22 with a considerable variance (*SD* = .28) among healthy controls when they had to detect whether a target utterance was a joke (involving irony) or a lie. So in this study some participants detected the irony in all items whereas some were not performing above chance level.

In one of the few studies assessing irony *production*, Matthews, Hancock, and Dunham (2006) found considerable interindividual variance (*SD*s): when their participants had to choose between a literal and an ironic communicative response to eight situations, on average around half of the criticisms (*M* = 4.31, *SD* = 1.14) and one fourth (*M* = 2.00, *SD* = 1.75) of the compliments were delivered ironically, that is, by mock compliments and mock criticisms, respectively. Hence, also the use of ironic utterances varies between individuals. In a comparable paradigm used by Ivanko et al. (2004) the resulting individual differences in statement selection in a forced choice production task were meaningful in terms of self-reported use of sarcasm explaining variance in the choices made (self-reported use of sarcasm was positively related to the frequency of choosing sarcastic statements). To summarize, there appears to be a pattern in the results reported in studies investigating irony detection performance. First, performance means are usually rather high (i.e., low error rates or high hit rates), pointing to a low difficulty of the tasks employed or selective sampling. Second, despite low difficulty (and possibly resulting ceiling effects) there is still variance that can be possibly made sense of in terms of explaining it by variables specific to the person. Furthermore, there is interindividual variance in irony production performance.

The question arising now is whether the individual difference variance can be fully explained by established traits, or whether there is something beyond. We argue that individual differences are linked to irony in two ways: First, the detection and production of verbal irony will be subject to established traits. Second, we expect individual differences in the extent to which people detect irony, produce irony, seek irony, or enjoy irony (i.e., the

sense of irony) over and above the variance shared with known variables. We will next present a set of hypotheses concerning those two notions.

Explaining differences in irony performance by known traits and ability

It can be assumed that established traits and abilities to a certain degree impact on the typical level of irony performance (i.e., in unobtrusive tasks), whereas general mental ability restricts the maximal level of performance (when being explicitly instructed to detect or produce irony). Involving cognitive and emotional processes, such as inferring another person's belief about the current state of affairs and inferring, identifying and understanding emotions (Ziv, Leiser, & Levine, 2011), irony detection can be argued to be subject to personality traits with cognitive and emotional components, such as biased beliefs or emotional responses. The production of irony may be subject to traits influencing interpersonal and humorous behavior, such as teasing, acting, and joking (e.g., Keltner, Capps, Kring, Young, & Heerey, 2001). As one of these traits, the histrionic self-presentation style was conceptualized to also go along with irony production (Renner et al., 2008). Histrionic self-presentation is defined as "a way of shaping everyday interactions by explicit As-If-behaviors" (Renner et al., 2008, p. 1303). As-If-behaviors, typically employed in order to "gain attention, entertain others, liven up a situation, create good mood and to relieve stress and tension in oneself and others" (Renner et al., 2008, p. 1305), also encompass verbal irony.

Once more bridging irony and humor, we would like to hypothesize a link between the mis-detection of verbal irony and the fear of being laughed at (gelotophobia, Ruch & Proyer, 2008). The fear of being laughed at is a personality trait characterized by the bias to experience a broad range of social interactions involving humor (which gelotophobes generally misperceived as put-down humor if directed at them) and laughter (generally perceived as victimizing) as hurtful attacks, ridicule and contempt (see Ruch, Hofmann, Platt & Proyer, 2014, for a review). There are several reasons why gelotophobes can be assumed to

be prone to misses but also false alarms in irony detection whereas non-gelotophobes do not. First, gelotophobes are characterized by having a strong sensitivity towards offense (Titze, 2009) and a low self-esteem (see Ruch et al., 2014). Being convinced to be deficient they can be considered to have a tendency to expect being criticized by others, especially in a derisive way, making them prone to suspect ironic compliments to be literal criticism. Secondly, given that "biting sarcasm" can be employed to ridicule others, gelotophobes may suspect irony when being addressed by compliments, because they have a bias to expect being ridiculed. Furthermore, irony can be used when putting down others in a derisive or sarcastic way. Hence, as a third possible explanation, gelotophobes may have experienced traumatizing events with ridicule conveyed by ironic compliments and thereupon –with a paranoid tendency to anticipated ridicule (Platt, Ruch, Hofmann, & Proyer, 2012)− view others as likely to address them with sarcasm. Hence, we assume that gelotophobia affects verbal irony detection, specifically when dealing with ironic compliments expressed via a mock criticism and also when dealing with literal compliments (in terms of a false positive irony detection).

Aside from gelotophobia and still staying in the realm of humor, there is reason to expect that the temperamental basis of humor affects irony performance. High-scorers in seriousness (as assessed with the State-Trait Cheerfulness Inventory–STCI, Ruch, Köhler, & van Thriel, 1996) are described by "the preference for a sober, object-oriented communication style (for example, saying exactly what one means without exaggeration or ironic/sarcastic undertones)" (Ruch et al., 1996, p. 308). Thus, serious individuals, preferring a *bona fide* communication mode, may be inclined to miss out on irony when expecting pragmatic communication. Furthermore, seriousness might go along with a reduced readiness to process play signals typically going along with ironic teasing (Keltner et al., 2001).

When looking at irony *production*, two different dispositions towards laughter and ridicule can be expected to be relevant: interindividual tendencies to either a) *enjoying being laughed at* (i.e., gelotophilia; Ruch & Proyer, 2009), or b) *enjoying to laugh at others*

(katagelasticism; Ruch & Proyer, 2009). While for example ironically criticizing oneself (by mock self-praise) might be a behavior aiming at making others laugh at one's expense, ironic criticisms directed at others may serve to expose the reasons why others can be laughed at. Furthermore, bad mood (assessed with the STCI, Ruch et al., 1996) and especially the facet *ill-humoredness* (sullen, grumpy, grouchy feelings; Ruch et al., 1996) can be reasoned to go along with the expression of a negative attitude, hence making ironic criticisms more likely.

General mental ability in terms of intelligence has been assessed in several of the studies comparing patients and healthy controls as to their irony detection performance. Gaudreau et al. (2013) for example report a substantial association between irony detection and executive functions (with 21% shared variance). Mitchley, Barber, Gray, Brooks, and Livingston (1998) found a comparable correlation between the rate of errors made when cognitively appraising the meaning of sarcastic utterances and a non-verbal measure of general intellectual ability (28% shared variance) but not with a measure of premorbid verbal intelligence among n =13 patients with schizophrenia. The authors explain the absence of a homologous relationship among healthy controls by the lack of variance in the irony performance task in this group (i.e., all controls answered all of the sarcastic items correctly). This implies that more difficult tasks are needed in order to avoid ceiling effects among healthy adults. Varga et al. (2014) report similar results when looking at the association between irony comprehension performance and general intelligence (27% shared variance). The strength of these relationships can be viewed as sufficiently high to demonstrate that there is meaningful variance in irony comprehension performance that relies also on intellectual ability, but sufficiently low to discriminate irony comprehension from general intellectual ability. As regards irony *production*, there is not more than a weak hint at the role of general mental ability with intelligence being slightly correlated (i.e., indicating 14% shared variance) to self-reported use of irony as assessed with one item (Milanowicz, 2013).

To conclude, our position is that there are meaningful differences in irony performance that we can make sense of in terms of explaining it by personality or general mental ability (intelligence). Although irony performance is argued to rely on processes and mechanisms that are subject to personality and intelligence, meaningful individual differences in irony performance may not be fully accounted for by established traits. This consideration leads us to hypothesize that there may be a trait that we label the *sense of irony*. Next to entailing an ability component (as assessed by the maximal performance), the *sense of irony* can be reasoned to have a habitual component (as assessed by the typical performance). Thus, we next look into how the "sense of irony" could be conceptualized and eventually measured.

What is the *sense of irony*?

We would like to bring to discussion that person-specific affective and cognitive processing patterns important for irony performance culminate in a trait that we label the *sense of irony*. Traits are relatively stable over time and consistent across situations. We hypothesize that when dealing with ironic utterances and situations eliciting the production of irony, a person's irony performance also depends on his or her *sense of irony*. For a more complete understanding of irony (and for successful experimenting or assessment) we want to distinguish among the following components: First, maximal irony behavior must be distinguished from typical irony behavior. Maximal irony behavior refers to a person's capacity to produce or detect irony. It determines the upper limit of possible performance in the sense of ability. Characteristically, the maximal irony behavior would be assessed with a performance test, just like intelligence. Typical irony behavior denotes a person's habitual level of irony production and detection. If not explicitly instructed to watch out for irony (or produce such; i.e., in an unobtrusive test), the rate at which individuals detect (or produce) irony would resemble this habitual component. For example, cultural rules, habits and expectations may explain why a person can score high in an explicit irony production test but at the same time hardly ever uses irony in real life. Apart from these factors, we argue that the

sense of irony encompasses enduring tendencies toward a certain level of irony performance in relation to the upper limit of one's full capacity. Furthermore, the *sense of irony* may predict the emergence of mind-sets facilitating or impeding irony detection and production. Therefore, individuals high in *sense of irony* may be inclined to get into a state (e.g., bad mood) where they are prone to produce irony more readily when joining a group and also have a higher readiness to detect playful signals cueing irony (facilitating irony detection).

Assessment of the *sense of irony*

There is a need to develop means of assessment that make variance in irony performance measurable. In general, tests are needed that allow for the assessment of habitual levels of irony performance (i.e., the typical behavior rather than only maximal behavior). This could be attained by unobtrusive tests or by means of utilizing more ambiguous stimuli.

A test for the assessment of the *sense of irony* needs to employ items that are more difficult than the performance tasks and ad-hoc tests developed so far. Item difficulties should show a range allowing for a differentiation between subjects on the whole spectrum of the variable. Also, the evaluation of non-ironic distractor items is essential in order to cover false positive detection of irony (cf. Kreuz, 2000). In terms of signal detection theory (Green & Swets, 1966), research on verbal irony detection has –to our knowledge without exception– focused on factors leading to false negative detection of verbal irony, neglecting the investigation of false positive detection; i.e., taking literal language for verbal irony (false alarms). We hypothesize that the presumed phenomenon of false positive irony detection is not limited only to individuals with the fear of being laughed at (gelotophobia). Also non-gelotophobes can be assumed to differ in their inclination to falsely detect irony in literal utterances. Given that situations involving literal language can be ambiguous and, therefore, misperceived, false positive irony detection can be considered to also have high relevance in social and professional functioning (such as joking, flirting, teasing, negotiating, debating,

etc.). Literal language is far more frequent than verbal irony, which according to Gibbs's (2000) findings can be estimated to occur with a frequency of 8% of turns taken (in conversations among friends). Hence, the possibilities provided for mistaking literal utterances as irony are by nature more frequent than instances when ironic language could remain unrecognized.

Outline of future directions and possible applications

To summarize, we argued that irony can be described to get frequently used with a humorous intent, both in benevolent and disparaging ways. Furthermore, we reviewed evidence stressing the need for considering interindividual differences in irony performance. Also, we recommend developing assessment tools for both, the typical and maximal irony behavior in its different facets (i.e., considering also false positive irony detection).

We propose that there are at least three ways to apply the assessment of individual differences in irony performance to new research questions. First, focusing on individual differences in irony performance is a prerequisite to study who is able (and inclined) to produce and experience the humor in irony (also in the disparaging part). Second, in the face of first evidence for regional differences in self-reported use of sarcasm (Dress, Kreuz, Link, & Caucci, 2008), new standardized tests can fuel cross-cultural research on irony. Third, correlates and mechanisms of irony performance can be studied in a more targeted (and more controlled) fashion if we develop tests tapping into interindividual variance in irony performance among healthy adults instead of resorting to subjects with disorders.

PART II

Studying irony detection beyond ironic criticism: Let's include ironic praise

Richard Bruntsch and Willibald Ruch

The official citation that should be used in referencing this material is:

Bruntsch, R., & Ruch, W. (2017). Studying irony detection beyond ironic criticism: Let's include ironic praise. *Frontiers in Psychology*, *8*:606. doi:10.3389/fpsyg.2017.00606

Abstract

Studies of irony detection have commonly used ironic criticisms (i.e., mock positive evaluation of negative circumstances) as stimulus materials. Another basic type of verbal irony, ironic praise (i.e., mock negative evaluation of positive circumstances) is largely absent from studies on individuals' aptitude to detect verbal irony. However, it can be argued that ironic praise needs to be considered in order to investigate the detection of irony in the variety of its facets. To explore whether the detection ironic praise has a benefit beyond ironic criticism, three studies were conducted. In Study 1, an instrument (Test of Verbal Irony Detection Aptitude; TOVIDA) was constructed and its factorial structure was tested using N = 311 subjects. The TOVIDA contains 26 scenario-based items and contains two scales for the detection of ironic criticism vs. ironic praise. To validate the measurement method, the two scales of the TOVIDA were experimentally evaluated with N = 154 subjects in Study 2. In Study 3, N = 183 subjects were tested to explore personality and ability correlates of the two TOVIDA scales. Results indicate that the co-variance between the ironic TOVIDA items was organized by two inter-correlated but distinct factors: one representing ironic praise detection aptitude and one representing ironic criticism detection aptitude. Experimental validation showed that the TOVIDA items truly contain irony and that item scores reflect irony detection. Trait bad mood and benevolent humor (as a facet of the sense of humor) were found as joint correlates for both ironic criticism and ironic praise detection scores. In contrast, intelligence, trait cheerfulness, and corrective humor were found as unique correlates of ironic praise detection scores, even when statistically controlling for the aptitude to detect ironic criticism. Our results indicate that the aptitude to detect ironic praise can be seen as distinct from the aptitude to detect ironic criticism. Generating unique variance in irony detection, ironic praise can be postulated as worthwhile to include in future studies—especially when studying the role of mental ability, personality, and humor in irony detection.

Keywords: Cheerfulness, Confirmatory Factor Analysis, Corrective humor, Intelligence, Ironic praise, Irony, Personality, STCI

Introduction

Ironic criticism and ironic praise can be distinguished as two basic types of verbal irony (cf. Kreuz & Link, 2002). The two types are structurally similar to each other as both involve mock evaluations of circumstances with a valence opposite to the speaker's true appraisal. As the characteristic difference between the two, ironic praise is characterized by a negative valence in what is said and a positive valence in the appraisal that is ironically implied, while in ironic criticism the converse is true.[8]

When we use irony, we typically utter something different from what we want to express, i.e., typically the opposite of our true appraisal of circumstances. Characteristically, we expect the listener to recognize our overt dissimulation by seeing through the counterfactual nature of our utterance and to eventually detect the intended meaning of what we say nonetheless (Groeben & Scheele, 2003). However, this is not always the case, as listeners may not detect the irony for certain reasons. For example, imperfect irony detection rates were found as a function of the ambiguity of the context of ironic utterances. Accordingly, Ackerman (1983) reports considerable average error rates in his irony detection task (ranging from 5.6% to 24.1% depending on the difficulty of the stimuli) in a control group consisting of college students. Furthermore, individuals differ in their aptitude to detect

[8] To illustrate: imagine that a circle of friends is watching a sports match and some of the attendees support Team A while other attendees support Team B. An example of ironic criticism would be if one of the supporters of Team A said "Terrific shot! You're handing us a resounding defeat!" when a player of Team B tries but fails to score a goal in the match (for example when the speaker wants to ridicule the arrogant prediction made by one of the supporters of Team B that "their" Team B would win at a canter). In contrast, if one of the supporters of Team A said "Terrible shot! We don't stand the slightest chance!" when a player of Team A scores a goal, this would be an example of ironic praise (for example when the speaker wants to ridicule one of the supporters of Team B for his or her arrogant prediction that Team A would lose the match in a sad spectacle of defeat).

verbal irony, which results in systematic variance in irony detection performance (e.g., Winner, Brownell, Happé, Blum, & Pincus, 1998; see Bruntsch, Hofmann, & Ruch, 2016, for an overview).

In the studies investigating irony detection, a plethora of tasks and *ad-hoc* test has been used to assess individuals' aptitude to detect verbal irony. However, most of these studies did not utilize both *ironic criticism* (as a mock positive evaluation of negative circumstances) and *ironic praise* (as a mock negative evaluation of positive circumstances). Rather, the stimuli used in the existing studies on irony detection mostly rely on ironic criticisms (such as in the form of sarcasm[9]), whereas ironic praise is not represented (e.g. Ackerman, 1983; Happé, 1993; McDonald & Pearce, 1996; Mitchley, Barber, Gray, Brooks, & Livingston, 1998). This is somewhat puzzling, as ironic praise can be found as counterbalanced with ironic criticism in the stimuli used in studies targeting different aspects of irony processing, such as when investigating processing times (i.e., response latencies) of ironic stimuli vs. their literal counterparts (Schwoebel, Dews, Winner, & Srinivas, 2000). Likewise, there are studies investigating perceived speaker's intent in "ironic insults" (matching the definition of *ironic criticism* we adhere to; cf. Kreuz & Link, 2002) and "ironic compliments" (matching the definition of *ironic praise*) vs. direct insults and direct compliments, respectively (e.g., in terms of ratings of mocking and politeness, i.e., Pexman & Olineck, 2002).

[9] The terms "irony" and "sarcasm" are sometimes used interchangeably (e.g., Pexman & Olineck, 2002) and there is an ongoing debate as to whether sarcasm and irony are essentially the same thing (cf. Attardo, 2000). However, we wish to adhere to a demarcation between irony and sarcasm in terms of two naturally overlapping but conceptually distinct phenomena. For example, in the *Merriam-Webster Dictionary* sarcasm is defined as "a sharp and often satirical or ironic utterance designed to cut or give pain". Irony can be seen as related to sarcasm because phrasing a criticism ironically was found to enhance the degree of perceived condemnation (as compared to phrasing it literally, i.e., Colston, 1997). In the present paper we will stick to the term *irony* (even if sarcasm is involved in a specific instance of irony), foremost because if any of the studies in our literature review use the term *sarcasm*, they originally refer to *ironic* sarcasm (rather than *non-ironic* sarcasm).

However, studies investigating irony *detection* have largely neglected the sampling of ironic praise stimuli. This may be owed to the view that ironic praise can be seen as the less prevalent and less "prototypically ironic" type of irony (cf. Kreuz & Link, 2002). However, a study by Langdon, Davies, and Coltheart (2002) demonstrated that stimuli containing ironic praise led to different results than ironic criticism stimuli. Langdon et al. (2002) used both, ironic criticism (labeled as *sarcasm*) and ironic praise (labeled as *banter*), and distinguished them in separate scores for their investigation of irony detection in schizophrenic patients vs. normally functioning control subjects.[10] As Langdon et al. (2002) report, ironic praise was harder to detect than ironic criticism, especially in the group of patients with schizophrenia. Thus, it can be hypothesized that ironic praise may be the very type of irony that is affected by impaired or unusual cognitive and affective functioning. More generally, it may be suggested that ironic praise leads to meaningful interindividual variance in irony detection tasks beyond the one found for ironic criticism.

Ironic criticism vs. ironic praise

As detailed below, we argue that the two types of irony can be distinguished considering at least three aspects: (a) they have different purposes and functions in communication, (b) in irony detection ironic praise may depend on individuals' expression of certain traits more

[10] Langdon et al. (2002) use the term *banter* when labeling the category of their stimuli containing a negative statement being used in a positive context (which corresponds to the definition of ironic praise we adhere to) without the intention to harm or to criticize, whereas their *sarcasm* stimuli were characterized by a positive statement used in a negative context (which corresponds to the definition of ironic criticism) with the intention to harm or to criticize. It is necessary to mention here that elsewhere bantering irony was conceptualized to occur not only in situations in which the speaker intends to ironically praise (i.e., in terms of *kind banter*) but also when *ironically criticizing* (i.e., in terms of *sarcastic banter*; cf. Anolli, Ciceri, & Infantino, 2002). However, as far as we can tell from their report, Langdon et al. (2002) used banter only in the form of *kind banter* (i.e., they did not include *sarcastic banter* involving ironic criticism) in their banter stimuli.

than ironic criticism, and (c) in irony detection they demand different cognitive and affective processes in individuals.

(A) One may characterize that ironic praise is typically used for different purposes (for example good-natured "ironic teasing"; Keltner, Capps, Kring, Young, & Heerey, 2001) than ironic criticism (for example aggressive ridicule). In the form of teasing, ironic praise may be reasoned to be a way to humorously apprise the recipient of social norms when harmless transgressions occur—such as when using it as a playful provocation in socializing, flirting, or entertaining. In contrast, ironic criticism may be employed for the purpose of apprising the recipient of social norms when more severe transgressions occur—such as when resolving conflicts by aggressive ridicule (cf. Norrick, 1994; Keltner et al., 2001). Furthermore, as ironic praise is typically used in the face of positive circumstances, one may reason that ironic praise is more suitable than ironic criticism (which in turn is typically used in the face of adverse circumstances) for certain of the discourse goals found for verbal irony, such as to be funny or witty, to be humorous, and to play or to be silly (cf. Kreuz, Long, & Church, 1991).

(B) The different functional aspects of the two types of irony (such as different utilities in social interaction) may affect the detection of ironic criticism and ironic praise differently, depending on individuals' expression of certain traits, such as the sense of humor. As the notion that humor is a function of irony is pervasive in the literature (cf. Bruntsch et al., 2016), the sense of humor (which can be defined as relatively stable interindividual differences in the tendency to react to humor and to produce humor, and a serene attitude toward life; see Ruch, 1998) can be assumed to go along with the readiness to detect or mis-detect verbal irony. Certain *facets* of the sense of humor may come into play more evidently in the detection of ironic praise than in the detection of ironic criticism. Furthermore, looking at ironic praise as a playful and light-hearted figure of speech, its detection may be facilitated

by cheerfulness (e.g., Ruch, Köhler, & van Thriel, 1996) more than this is the case for ironic criticism. This may be the case because highly cheerful individuals may process cues signaling playfulness more readily, which helps to reject the uttered negative evaluation and detect the more positive implication of ironic praise. Importantly, this may not hold true for ironic criticism, which may be seen as less playful and less jocular than ironic praise.

(C) It can be argued that the norm violation that irony typically alludes to and criticizes (e.g., Garmendia, 2014; Utsumi, 2000) is harder to recognize in the case of ironic praise: it may be more obvious and hence easier to understand why ironic criticism is used. This may be because people generally have positive expectations (e.g., successful players in professional sports; cf. Kreuz & Link, 2002). Thus, the detection of ironic praise may require a more complex mental representation of the background of the ironic remark and a more effortful cognitive search for the antecedent event that ironic remarks typically refer to (Kreuz & Glucksberg, 1989), as compared to the detection of ironic criticism. In line with this consideration, intelligence may be more relevant for the detection of ironic praise than for the detection of ironic criticism. If the role of intelligence truly was more evident in the detection of ironic praise, ironic praise should be included in irony research when mental abilities as well as mental impairments are targeted.

Aims of the paper

The current paper has three main aims. Firstly, a test for the assessment of irony detection with two different scales (i.e., ironic criticism vs. ironic praise) will be developed, opting for an indirect measurement format (Study 1). It is aimed to use two testing modes with different degrees of *irony alertness*: hiding the measurement intention from participants (i.e., *irony non-alert* mode) vs. making irony salient (*irony alert* mode). Using confirmatory factor analysis, the two-factor structure (corresponding to the distinction between ironic criticism

and ironic praise) will be tested. Secondly, in Study 2 we will validate the soundness of the stimuli and the indirect measurement by (a) using an experimental approach (i.e., comparing four testing conditions: irony alert testing, irony non-alert testing, forced ironic interpretation, and forced literal interpretation), (b) testing whether there is a convergence between the test scores and direct irony-ratings, and (c) comparing direct irony-ratings between ironic items and non-ironic distractor items (which should differ from each other). Thirdly, Study 3 will explore ability and personality correlates of the two scales. It is expected that ironic praise detection scores are at least as strongly related—if not even more strongly related—to (a) intelligence, (b) the ability to distinguish irony from a lie, (c) different facets of the sense of humor, and (d) traits constituting the temperamental foundation of the sense of humor (e.g. cheerfulness), as this is the case for the detection of ironic criticism.

Study 1: Development of the Test of Verbal Irony Detection Aptitude (TOVIDA)

It is assumed that there is meaningful interindividual variance in irony detection performance in terms of an irony detection aptitude. It is hypothesized that this aptitude comprises two facets: the aptitude to detect ironic criticism and the aptitude to detect ironic praise. After selecting those items with the most acceptable psychometric features, a confirmatory factor analysis will be employed to investigate whether the two predefined concepts used in the instrument (ironic criticism and ironic praise) are represented by two different structural components. A first sample will be used to determine psychometric properties under *irony non-alert* testing conditions, as this unobtrusive method can be reasoned to reflect individuals' everyday mode of dealing with irony (i.e., usually, we do not deliberately watch out for irony). Then, a second sample will be used for cross-validation to see whether the fit of a two-factor model (i.e., ironic criticism vs. ironic praise) can be confirmed under *irony alert* testing conditions. Maximizing irony alertness can be reasoned to reduce systematic

noise in the interindividual variance. To specify: as some individuals may be more biased not to anticipate irony in a psychological survey than others, *irony non-alert* testing presumably would lead to artificial co-variance between the items. Furthermore, as the shared variance between items systematically depends on the interindividual variance that makes co-variance arise in the first place, this method can be seen as a source of data accommodating a more conservative test of the assumed model.

Methods

Participants

Participants were recruited via university mailing lists, social platforms, and leaflets. Two independent samples were used. Sample 1 consisted of 152 German-speaking subjects (40 males [35.7%]). Age in Sample 1 ranged from 18 to 51 years with a mean of 22.8 (*SD* = 5.8). Sample 2 consisted of 159 German-speaking subjects (39 males [32.5%]). Age in Sample 2 ranged from 18 to 67 years with a mean of 24.1 (*SD* = 7.3).

Materials

Test of Verbal Irony Detection Aptitude-40 (*TOVIDA-40*). To develop a test for the assessment of irony detection aptitude, thirty scenarios containing ironic target utterances (among which 20 contained ironic criticism and ten contained ironic praise) and ten scenarios with non-ironic target utterances were written using a rational construction procedure. Irony detection was defined as the comprehension of the true meaning of ironic target utterances as opposite to the literal meaning in ambiguous situations short of distinct information. Each scenario consists of a short story about two or more people and culminates in a final utterance (the target utterance) made by one of the protagonists. Target utterances contain either verbal irony or literal speech. When generating the stimuli, irony was designed as follows: in the ironic utterances used in the *ironic criticism* stimuli, speakers (i.e., the story characters making the target utterance) use a choice of words which, when used non-ironically, denotes

a positive appraisal—while ironically implying an opposite (i.e., negative) appraisal. Conversely, as the characteristic feature of the utterances found in the *ironic praise* stimuli, speakers use a choice of words which, when used non-ironically, denote a negative appraisal of circumstances—while ironically implying an opposite (i.e., positive) appraisal. In the ironic criticism stimuli, speakers comment on a negative circumstance described in the short story (with a mock positive evaluation). In contrast, in the ironic praise stimuli, speakers comment on a positive circumstance (with a mock negative evaluation). In the TOVIDA-40, ironic utterances typically involve meta-messages indirectly implied by the speaker, such as when mocking the addressee's overly self-critical or self-effacing attitude.[11]

The scenarios are designed as ambiguous in order to warrant sufficient psychometric item difficulty, i.e., to avoid ceiling effects. This is why the stories still make some sense when irony is not detected in the ironic items (i.e., in the case of false negative detection) and when irony is falsely detected in the non-ironic items (i.e., in the case of false positive detection). Accounting for ambiguity in the process of irony detection, Utsumi (2000) points out that irony is distinguished from non-irony by assessing the degree to which a given utterance resembles prototypical irony. That is, not every ironic utterance unambiguously fulfills the constituting criteria of irony. Rather, the listener detects irony by assessing the similarity between a given utterance and a prototype of irony. Hence, ambiguity can be seen as a typical feature of real-life situations involving irony. However, in the scenarios of the TOVIDA-40 there are unobtrusive cues signaling the preconditions for the ironic utterance, i.e., hints to a reason for the speaker to express a negative attitude via ironic criticism or

[11] As speakers say something different from what they actually want to express, irony classifies as an indirect speech act, cf. Holtgraves, 1997). What makes irony different from other forms of indirect speech acts is that ironic speech acts are characterized by an overt insincerity. That is, ironic speakers achieve indirectness by engaging in an evident dissimulation when inversing the valence of their true appraisal in the verbatim utterance (i.e., especially by using a choice of words denoting *the opposite* of their true appraisal of circumstances, cf. Attardo, 2000, 2001b).

ironic praise (cf. Garmendia, 2014; Utsumi, 2000)[12]. In order to assess whether participants chose a literal or an ironic interpretation of target utterances, participants have to judge scenarios along statements about factual aspects of the situation or actors' emotional states as causes or consequences of target utterances. A person detecting the irony correctly appraises the situation differently from a person not detecting the irony. The TOVIDA-40 was designed as an unobtrusive test that can be optionally administered without any mention of irony and distracts test-takers from its true measurement intention. Six statements are provided for the appraisal of the situation (to be rated on a four-point scale ranging from 1 = "does not apply at all" to 4 = "fully applies"), among which three are indicative of irony detection (see Appendix A2). The other three appraisal statements are designed to distract from the intention of the task. For example, there is a statement asking whether the protagonists behave like a typical male or female (according to his or her gender) provided for every scenario. A high item score in the ironic items indicates correct positive irony detection, i.e. the comprehension of the true meaning of ironic target utterances as opposite to the literal meaning. The ironic items are administered alternating with the non-ironic distractor items.

Procedure

Participants were tested individually using an online-survey. They were randomly assigned to one of two groups (labeled here as Sample 1 and Sample 2). They either were instructed without any mention of irony (Sample 1: *irony non-alert* testing) or provided with a definition of verbal irony and instructed to watch out for irony in the stimuli, i.e., they were

[12] As Garmendia (2014) argues, irony is always negative in terms of a critical attitude. That is, also in the case of ironic praise, which—as a meta-message—can be described to typically involve a hint to the transgression of (sometimes unwritten) rules, for example the norm of not to be vain, not to boast, not to be arrogant, not to be overly modest, not to make false promises, and so on. That does not mean that there is not another meta-message on a higher level that can be characterized as benevolent and more positive. For example, ironic teasing can be corrective and bonding at the same time, as the teaser implies that he or she thinks that the relationship with the teased person is strong and close enough to make playful provocation possible without risking a serious social damage (cf. Boxer & Cortés-Conde, 1997; Norrick, 1994).

instructed that some of the scenarios they were about to appraise contain verbal irony whereas others do not (Sample 2: *irony alert* testing). Participants completed the TOVIDA-40 after they filled in questions about their demographic features and German language proficiency.

Preliminary Analyses

In order to arrive at more reliable items scores, two of the three indicative statements were selected for every item applying a scale reliability criterion: inter-correlations between the three indicators were computed using Sample 1. The two indicators with the highest inter-correlation were selected and averaged to generate the item scores. In order to attain a more economic form of the TOVIDA-40, corrected item-total correlations (CITCs) were computed and considered as a selection criterion. Ironic criticism and ironic praise items were analyzed separately in this step. For selection purposes, only Sample 1 was used. For each of the two sub-scales eight items showed CITCs of $r_{cit} \geq .45$ and were selected to build two scales to be analyzed in the further steps of Study 1. The sixteen selected ironic items and the ten non-ironic distractor items taken from the TOVIDA-40 will be referred to as the *TOVIDA* in the following sections.

Results

Internal consistencies of the two resulting sub-scales were sufficiently high. Cronbach's alpha was .83 (.76) for the ironic criticism scale and .83 (.77) for ironic praise scale in Sample 1 and Sample 2 (values for Sample 2 in brackets).

Within the irony alert sample (Sample 2), the fit of two different structural equation models was estimated. In the *assumed model*, two inter-correlating factors were modeled: one factor was defined by ironic criticism items and the other factor by ironic praise items. In the *control model*, a single factor was modeled defined by both ironic criticism and ironic praise items. As it turned out, the assumed two-component model had acceptable fit ($\chi^2 = 153.296$,

df = 103; Bentler Comparative Fit Index [CFI] = .906; root mean square error of approximation [RMSEA] = .056 [90% CI: .036; .073]; standardized root mean square residual [SRMR] = .0643). In contrast, the control model did not show acceptable model fit (χ^2 = 227.025, *df* = 104; CFI = .771; RMSEA= .078 [90% CI: .071; .102]; SRMR = .0840). The path coefficients for the assumed two-factor model are given in Figure 1. As Figure 1 shows, the ironic criticism scale and the ironic praise factors were substantially intercorrelated.

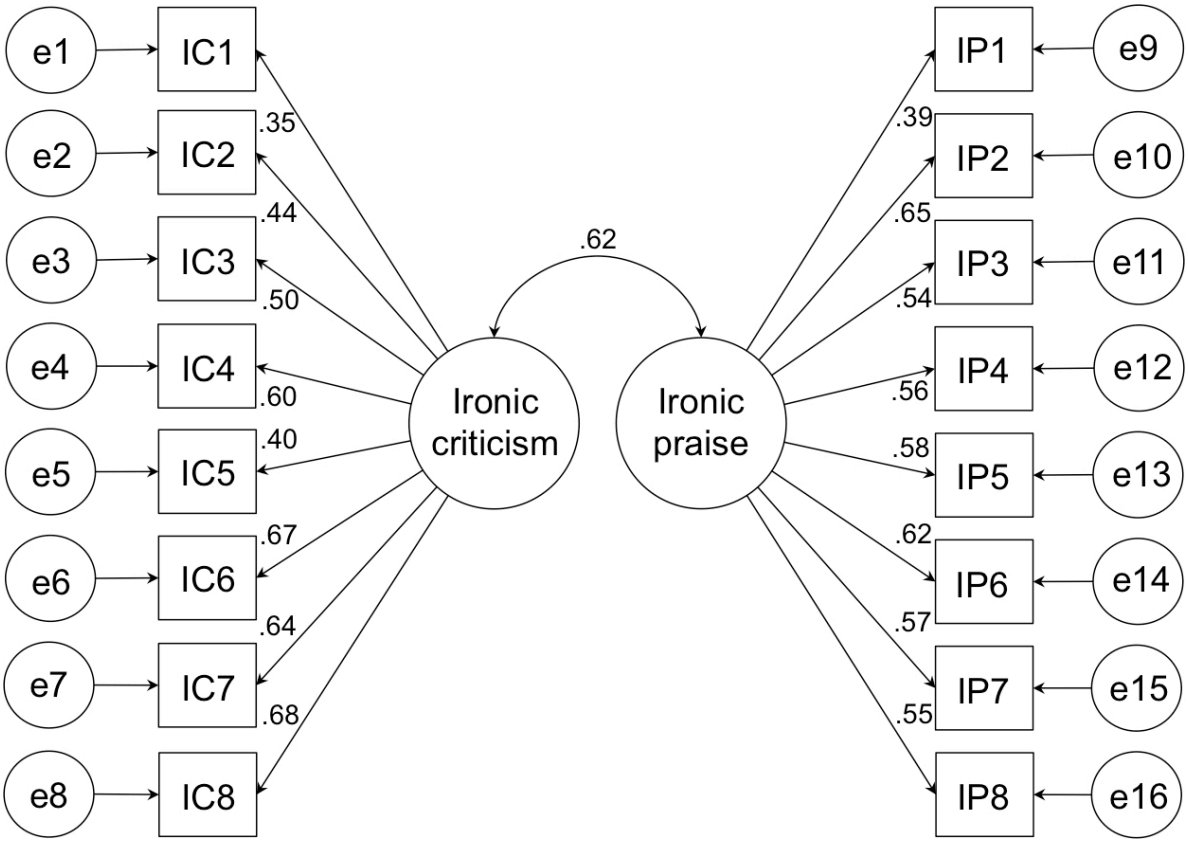


Figure 1. Estimates for path coefficients in the two-factor model that was confirmed in Study 1.

Discussion

The selection from the two types of items resulted in two scales with sufficient internal consistency. This indicates that there is an underlying irony detection aptitude creating shared variance in the items. Furthermore, the two-factorial structure could be affirmed, implying

that ironic praise generated unique variance in the TOVIDA. Hence, the findings of Study 1 support the assumption that the aptitude to detect ironic praise is worth distinguishing from the aptitude to detect ironic criticism.

Study 2: Experimental evaluation of the TOVIDA

The stimuli employed in the TOVIDA were designed as ambiguous in order to warrant sufficient psychometric item difficulty, i.e., to avoid ceiling effects. Furthermore, irony detection is assessed indirectly in order to make a testing mode feasible in which subjects are non-alert to the occurrence of irony in the stimuli. So it was deemed necessary to validate that the stimuli of the TOVIDA truly contain irony, and if so, that high (vs. low) test scores truly indicate high (vs. low) irony detection performance. The aim of study 2 was to address these questions.

Four criteria were defined to evaluate whether the TOVIDA allows for the assessment of irony detection: firstly, participants in the *irony alert* group are expected to have higher scores than participants in a *forced literal appraisal* group (i.e., participants instructed to view all items as non-ironic). This criterion reflects the consideration that there must be a group consensus among participants who know about the intention of the test that differs from a forced appraisal opposite to the designed ironic content. Secondly, participants in the *irony alert* group are expected to have higher scores than the ones in the *irony non-alert* group. The rationale of this criterion is that irony detection is facilitated when participants are instructed to watch out for irony (vs. being not informed about the possible occurrence of irony). Thirdly, a *forced ironic appraisal* group (i.e., participants instructed to view all items as ironic) is expected to have higher scores than the *forced literal appraisal* group. This criterion aims at ensuring that the appraisals used for the indirect measurement (and hence the item scores) are sensitive to irony detection. As a fourth criterion, the item scores within

the *irony alert* group are expected to be positively correlated with direct appraisals (i.e., explicit ratings) of ironic content (these were assessed only in this group).

Methods

Participants

Participants were recruited in university lectures, via university mailing lists, social platforms, and leaflets. The sample consisted of 154 German-speaking subjects (26 male [16.9%]). Participants' age ranged from 18 to 56 years with a mean of 24.8 years (SD = 7.8). They were randomly assigned to one of four testing conditions and the groups did not differ significantly as to age (F[3, 150] = 1.69, p = .17), nor gender (F[3, 150] = 0.085, p = .97).

Instruments

The *Test of Verbal Irony Detection Aptitude* (*TOVIDA; see* Study 1 for description and Appendix A2 for an example item). Item scores were computed following the method of Study 1.

Procedure

In an online-survey, participants were randomly assigned to one of four test conditions: (1) one group was given a definition of verbal irony, was briefed that some of the scenarios they were about to see contain verbal irony whereas others do not, and instructed to take all target utterances as *ironic* when appraising the scenarios along the predefined statements (*forced ironic appraisal*), (2) one group was given a definition of verbal irony, was briefed that some of the scenarios they were about to see contain verbal irony whereas others do not, and instructed to take all target utterances as *literal* while appraising the scenarios (*forced literal appraisal*), (3) another group was given a definition of verbal irony, was briefed that some of the scenarios they were about to see contain verbal irony whereas others do not, and instructed to watch out for irony when appraising the scenarios according to their own interpretation (*irony alert*), and (4) the last group was instructed to appraise the scenarios

according to their own interpretation without any mention of irony (*irony non-alert*). More specifically, the experimental instructions in the *forced ironic appraisal* group and the *forced literal appraisal* group briefed participants (a) to willfully view the last sentence in each of the situations as ironic or non-ironic, respectively, and (b) to respond to all of the concerned questions *as if* the last sentence was truly ironic or non-ironic, respectively. In the irony alert group, participants were requested to make direct appraisals (i.e., explicit ratings) of ironic content in addition to the standard appraisal. These explicit ratings of ironic content were assessed via a four-point Likert-type scale (1 = "not ironic, 2 = "rather not ironic", 3 = "rather ironic", 4 = "ironic"), accounting for the ambiguous nature of the scenarios. Participants in the alert group were considered lay judges for this purpose (for the use of laypersons for validation purposes see Legree, 1995). The *irony alert* group was randomly over-sampled in order to warrant sufficient sample size for the planned correlational analyses.

Results

Do the stimuli of the TOVIDA contain irony?

Group means of item scores are given in Table 2. As Table 2 shows, all items met the criterion to verify that they contain irony. More precisely, in line with the expectations, the *forced literal appraisal* group had lower means than the *irony alert* group with medium to large effect sizes, indicating that generally irony is detected in ironic items. Furthermore, in the *irony alert* group item scores were generally higher than in the *irony non-alert* group with small to large effect sizes (however, only ten out of 16 of the comparisons yielded significant differences). In line with the expectation, being alert to irony facilitated irony detection.

Table 2

Descriptive and Test Statistics of Group Scores in Study 2.

Item	Test instruction (group)								Group comparisons					
	1 Non-ironic		2 Ironic		3 Alert		4 Non-alert		1 vs. 2		1 vs. 3		3 vs. 4	
	M	SD	M	SD	M	SD	M	SD	*t*(63)	*d*	*t*(90)	*d*	*t*(87)	*d*
IC1	2.17	0.57	3.18	0.54	2.54	0.67	2.47	0.72	-7.25*	1.83	-2.59*	0.58	0.45	0.10
IC2	1.60	0.70	3.25	0.62	3.04	0.70	2.76	0.89	-10.09*	2.52	-9.06*	2.06	1.55	0.37
IC3	1.36	0.42	3.23	0.86	2.82	0.88	2.00	0.89	-10.63*	2.65	-8.34*	1.89	3.92*	0.93
IC4	1.37	0.55	3.50	0.72	2.64	1.05	2.13	1.06	-13.11*	3.26	-6.06*	1.37	2.05*	0.48
IC5	1.90	0.64	2.88	0.59	2.46	0.57	2.35	0.75	-6.35*	1.60	-4.19*	0.95	0.80	0.18
IC6	1.69	0.67	3.36	0.64	3.02	0.68	2.67	0.94	-10.23*	2.56	-8.62*	1.96	1.95	0.46

Table 2 continues.

Table 2 continued.

Item	Test instruction (group)								Group comparisons					
	1 Non-ironic		2 Ironic		3 Alert		4 Non-alert		1 vs. 2		1 vs. 3		3 vs. 4	
	M	SD	M	SD	M	SD	M	SD	*t*(63)	*d*	*t*(90)	*d*	*t*(87)	*d*
IC7	1.56	0.74	3.49	0.72	3.07	0.80	2.16	1.04	-10.52*	2.65	-8.54*	1.93	4.42*	1.04
IC8	2.08	0.81	3.25	0.67	2.89	0.56	2.43	0.82	-6.37*	1.60	-5.48*	1.26	3.03*	0.72
IP1	2.21	0.55	3.71	0.49	3.52	0.58	2.87	0.69	-11.59*	2.90	-10.13*	2.29	4.51*	1.06
IP2	1.57	0.84	3.31	0.77	3.18	0.61	2.67	0.98	-8.66*	2.17	-10.32*	2.34	2.96*	0.70
IP3	1.77	0.57	3.34	0.78	2.77	0.69	2.48	0.69	-8.93*	2.25	-6.66*	1.52	1.78	0.42
IP4	1.38	0.48	3.34	0.93	2.95	0.88	2.38	1.01	-10.18*	2.55	-8.88*	2.01	2.61*	0.62

Table 2 continues.

Table 2 continued.

Item	Test instruction (group)								Group comparisons					
	1		2		3		4							
	Non-ironic		Ironic		Alert		Non-alert		1 vs. 2		1 vs. 3		3 vs. 4	
	M	SD	M	SD	M	SD	M	SD	*t*(63)	*d*	*t*(90)	*d*	*t*(87)	*d*
IP5	2.01	0.52	3.05	0.69	2.94	0.51	2.57	0.54	-6.65*	1.67	-7.99*	1.81	3.02*	0.71
IP6	1.90	0.67	3.43	0.52	3.22	0.55	3.01	0.78	-10.39*	2.60	-9.86*	2.24	1.40	0.34
IP7	1.44	0.44	3.33	0.84	2.83	0.86	2.15	0.82	-10.87*	2.71	-8.09*	1.83	3.43*	0.80
IP8	1.77	0.67	3.00	0.91	2.60	0.70	2.21	0.57	-6.03*	1.51	-5.28*	1.20	2.47*	0.59

Note. IC1-IC8 = ironic criticism items; IP1-IP8 = ironic praise items. Non-ironic = forced literal appraisal (n = 28), Ironic = forced ironic appraisal (n = 37), Alert = irony alert testing (n = 64), Non-alert = irony non-alert testing (n = 25). d = Cohen's d coefficient of effect size.

* $p < .05$.

Next, the direct appraisals of ironic content were examined to find out whether ironic items are viewed as more ironic than the non-ironic items. The frequencies of the single ratings were considered, given in Table 3. As Table 3 shows, ironic criticism items and the ironic praise items had numerically higher appraisals of being ironic ("rather ironic" and "ironic" answers) than non-ironic control items. It is noteworthy that the distributions of the proportions of ironic appraisals had a contact point: the ironic item with the lowest frequency of ironic appraisals (IC1) was judged about just as ironic as the non-ironic control item with the highest frequency of ironic appraisals (NC08). However, these two items can be seen as outliers in their group and as there was still a fair amount of judges consenting that the ironic items in question contain irony. Thus, they can be considered as difficult items but still containing irony. To test whether ironic criticism and ironic praise items were rated as more ironic than the non-ironic control items in the direct appraisals of ironic content, a mean of ratings over the eight items per scale was computed as well as the mean of ratings for the ten non-ironic control items. These scores were compared with paired sample *t*-tests. It turned out that the non-ironic control items were rated as less ironic ($M = 1.59$, $SD = 0.38$) than the ironic criticism items ($M = 2.95$, $SD = 0.58$, $t[63] = -14.47$, $p < .001$) and the ironic praise items ($M = 3.23$, $SD = 0.54$, $t[63] = -18.16$, $p < .001$), indicating large effect sizes (i.e., $d = 2.77$ and $d = 3.51$, respectively).[13]

[13] An exploratory analysis indicated that ironic praise items were appraised as somewhat more ironic than the ironic criticism items with a medium effect size ($d = 0.50$). This is important to point out, as the direct appraisals are substantially correlated with the item scores (i.e., the indirect appraisals). The irony in ironic praise items hence can be seen as less difficult to detect than ironic criticism. It is not clear whether this is owed to the fact that the ironic praise items used in the present set of studies are less ambiguous than the ironic criticism items or whether ironic praise *per se* is easier to detect than ironic criticism.

Table 3

Direct Irony Appraisal using Explicit Irony Ratings for the Single Items of the TOVIDA (Study 2).

Items	Rating scale steps							
	"not ironic"		"rather not ironic"		"rather ironic"		"ironic"	
	f	(%)	*f*	(%)	*f*	(%)	*f*	(%)
IC1	19	(29.7)	19	(29.7)	12	(18.8)	14	(21.9)
IC2	3	(4.7)	6	(9.4)	22	(34.4)	33	(51.6)
IC3	6	(9.4)	17	(26.6)	15	(23.4)	26	(40.6)
IC4	15	(23.4)	10	(15.6)	17	(26.6)	22	(34.4)
IC5	10	(15.6)	16	(25.0)	20	(31.3)	18	(28.1)
IC6	7	(10.9)	3	(4.7)	25	(39.1)	29	(45.3)
IC7	6	(9.4)	8	(12.5)	17	(26.6)	33	(51.6)
IC8	5	(7.8)	13	(20.3)	14	(21.9)	32	(50.0)
IP1	5	(7.8)	6	(9.4)	8	(12.5)	45	(70.3)
IP2	3	(4.7)	2	(3.1)	4	(6.3)	55	(85.9)
IP3	10	(15.6)	20	(31.3)	16	(25.0)	18	(28.1)

Table 3 continues.

Table 3 continued.

Items	Rating scale steps							
	"not ironic"		"rather not ironic"		"rather ironic"		"ironic"	
	f	(%)	*f*	(%)	*f*	(%)	*f*	(%)
IP4	7	(10.9)	5	(7.8)	23	(35.9)	29	(45.3)
IP5	2	(3.1)	3	(4.7)	25	(39.1)	34	(53.1)
IP6	2	(3.1)	3	(4.7)	13	(20.3)	46	(71.9)
IP7	8	(12.5)	12	(18.8)	14	(21.9)	30	(46.9)
IP8	11	(17.2)	12	(18.8)	23	(35.9)	18	(28.1)
NC01	33	(51.6)	21	(32.8)	7	(10.9)	3	(4.7)
NC02	49	(76.6)	14	(21.9)	1	(1.6)	0	(0)
NC03	34	(53.1)	18	(28.1)	4	(6.3)	8	(12.5)
NC04	37	(57.8)	20	(31.3)	2	(3.1)	5	(7.8)
NC05	47	(73.4)	9	(14.1)	6	(9.4)	2	(3.1)
NC06	34	(53.1)	23	(35.9)	6	(9.4)	1	(1.6)
NC07	56	(87.5)	6	(9.4)	1	(1.6)	1	(1.6)
NC08	18	(28.1)	20	(31.3)	13	(20.3)	13	(20.3)

Table 3 continues.

Table 3 continued.

Items	Rating scale steps							
	"not ironic"		"rather not ironic"		"rather ironic"		"ironic"	
	f	(%)	*f*	(%)	*f*	(%)	*f*	(%)
NC09	46	(71.9)	15	(23.4)	2	(3.1)	1	(1.6)
NC10	34	(53.1)	19	(29.7)	8	(12.5)	3	(4.7)
Descriptive statistics								
M_{IC}	8.88	13.9	11.50	18.0	17.75	27.7	25.88	40.4
SD_{IC}	5.49	8.6	5.68	8.9	4.33	6.8	7.24	11.3
M_{IP}	6.00	9.4	7.88	12.3	15.75	24.6	34.38	53.7
SD_{IP}	3.55	5.5	6.27	9.8	7.55	11.8	13.41	20.9
M_{NC}	38.80	60.6	16.50	25.8	5.00	7.8	3.70	5.8
SD_{NC}	10.84	16.9	5.48	8.6	3.80	5.9	4.03	6.3

Note. N = 64. IC1-IC8 = ironic criticism items, IP1-IP8 = ironic praise items, NC01-NC10 = non-ironic control items. M_{IC} / SD_{IC} = mean / standard deviation for ironic criticism item ratings; M_{IP} / SD_{IP} = mean / standard deviation for ironic praise item ratings; M_{NC} / SD_{NC} = mean / standard deviation for non-ironic control item ratings.

Is irony detection reflected in the item scores of the TOVIDA?

As Table 2 shows, the item score means of the *forced ironic appraisal* group were higher than item score means of the *forced literal appraisal* group, with large effect sizes. This indicates that a person will score high in all items if he or she detects the irony and score low if this is not the case. Finally, as expected, the direct appraisals (i.e., explicit ratings) of ironic content in the *irony alert* group correlated significantly with the respective item scores in all items with a mean of $r(63) = .72$, indicating good convergence between direct and indirect appraisals. This finding indicates that the TOVIDA test scores reflect the degree to which participants considered the stimuli as ironic.

Discussion

The results support the claim that, the ironic criticism and ironic praise stimuli used by the TOVIDA contain irony. Firstly, item scores were higher the group instructed to watch out for irony (i.e., the *irony alert* group) than in the group with experimentally induced minimal irony detection (i.e., in the *forced literal appraisal* group). This finding indicates that irony can generally be detected in the items of the TOVIDA (with a fair amount of interindividual variance, as shown by substantial standard deviations in *irony alert* and *irony-non alert* individuals' detection scores). Secondly, alertness to the ironic content of the stimuli fostered irony detection as the *irony-alert* group had higher item scores than the *irony non-alert* group in the majority of the items. Thirdly, the direct appraisals of the ironic content indicate that the ironic items were viewed as more ironic than the non-ironic items. There is also support for the claim that test scores reflect iron detection. Firstly, this was evident in terms of considerable differences between a group with experimentally induced minimal irony detection (i.e. in the *forced literal appraisal* group) and a group with experimentally induced maximal irony detection (i.e., in the *forced ironic appraisal* group). Secondly, the item scores corresponded well with direct appraisals (i.e., explicit ratings) of

ironic content. These findings indicate that the items of the TOVIDA assess irony detection performance and that the stimuli—although they were designed as ambiguous—were consented as containing verbal irony to an acceptable degree.

Study 3: Exploring the usefulness of ironic praise in a study of irony detection correlates

Study 3 aimed at exploring whether ironic praise stimuli have a benefit in the investigation of ability and personality correlates of irony detection. Among the preexisting studies assuming an individual differences perspective in irony research, Ivanko, Pexman, and Olineck (2004) explored the possibility to explain interindividual variance in an irony interpretation task (i.e., in terms of participants' ratings of speaker's intent, such as *sarcasm*, *mocking*, and *politeness*) by means of participants' scores in "conversational indirectness" (i.e., the tendency to phrase one's remarks indirectly and the extent to which a person looks for indirect meanings in the remarks of others, cf. Holtgraves, 1997). The present study aims to extend this and other previous work (e.g., Blouin & McKelvie, 2012) by (a) looking at irony *detection* (rather than irony *comprehension* as the interpretation of speaker's attributes in ironic utterances) and (b) including intelligence and a broad range of personality traits as individual differences variables.

As one of the hypothesized correlates, it may be argued that trait cheerfulness has a relevance especially to the detection of ironic praise as cheerful individuals may have a more positive outlook on themselves and others and hence be more inclined to expect jolly and jovial interactions involving playful ironic teasing rather than hostile and negative interaction involving serious ridicule, such as in the form of ironic criticism. Furthermore, certain facets of the sense of humor may be more relevant to the detection of ironic praise than to the detection of ironic criticism. According to Ruch and Heintz (2016a), the sense of humor

includes also two virtue-related facets, i.e., *benevolent humor* and *corrective humor*. As an accepting way of dealing with negative circumstances (e.g., human weaknesses), benevolent humor may be relevant especially to ironic criticism (typically occurring in the face of negative circumstances) but not as relevant to ironic praise (typically occurring in the face of positive circumstances). That is, individuals prone to use, enjoy, seek, and understand benevolent humor may have a higher aptitude to detect ironic criticism. The other facet is characterized by tendencies to wittily ridicule those who deserve it from a moral stance in terms of *corrective humor*. Importantly, irony is listed as one of the ways in which corrective humor manifests itself in speech. It can be argued that by exposing transgressions of social rules in a witty and playful way, corrective humor is conceptually more related to ironic praise than to ironic criticism, which in turn can be seen as the more serious and less ingenious form of irony. Hence, individuals who are prone to use, enjoy, seek, and understand corrective humor, may have a higher readiness to detect irony in the case of ironic praise more than in the case of ironic criticism.

Furthermore, irony detection can be related to mental abilities—and presumably especially so in the case of ironic praise. According to previous studies (e.g., Mitchley et al., 1998) intelligence can be seen as a prerequisite for the detection of ironic criticism. Under the presupposition that the detection of ironic praise poses a different cognitive challenge to the individual than the detection of ironic criticism, there may be a unique relationship between the detection of ironic praise and mental abilities. Hence, a test for the assessment of general mental ability (i.e., intelligence) will be employed. To include a measure of an ability more specific to irony detection, a task by Winner et al. (1998) will be jointly administered that was designed to assess the ability to discriminate between irony and lies among patients with brain damage. Simultaneously, by testing its convergence with the detection of ironic criticism and ironic praise, the convergent validity of the TOVIDA will be explored.

Accordingly, we expect that there are associations between ironic praise detection and individual differences variables that are robust beyond the influence of the variance the detection of ironic praise shares with the detection of ironic criticism. Moreover, it is expected that both of the two scales of the TOVIDA correlate positively with the irony/lie discrimination task, as the ability to distinguish irony from a lie can be seen as relevant to ironic praise to the same extent as to ironic criticism.

As a secondary aim, the association between the two scales of the TOVIDA and the Big Five personality traits will be explored to learn more about the discriminant value of the irony detection measure. It is expected that the Big Five as broad personality dimensions distal to the sense of humor and distinct from mental ability are largely unrelated to irony detection scores. For exploratory purposes, again two testing modes will be employed with different degrees of *irony alertness*: hiding the measurement intention from participants (i.e., *irony non-alert* mode) vs. making irony salient (*irony alert* mode). As there are no comparable previous studies on personality and ability correlates of irony detection, it was preferred to include both the irony non-alert and the irony alert mode of testing in order to safeguard the investigation against a selective method bias.

Methods

Participants

Participants were recruited in university lectures, and by means of university mailing lists, social platforms, and leaflets. Two independent quasi-experimental groups were tested. The first group (irony non-alert testing mode) consisted of 103 German-speaking subjects (28 male [22.0%). Age in Group 1 ranged from 18 to 38 years with a mean of 21.6 (*SD* = 3.5). Group 2 (irony alert testing mode) consisted of 80 German-speaking subjects (16 males [17.6%]). Age in this group ranged from 18 to 46 years with a mean of 22.7 (*SD* = 5.5).

Instruments

The *Test Of Verbal Irony Detection Aptitude* (*TOVIDA*) was used for the assessment of irony detection performance (see Study 1 for description/Appendix A2). Item scores were computed following the method of Study 1. The scores of the eight ironic criticism items and the eight ironic praise items were averaged to build an ironic criticism detection score and an ironic praise detection score, respectively. The internal consistencies of the two scales were comparable to those found in Study 1. Cronbach's alpha was .81 (.74) for the ironic criticism scale and .83 (.79) for the ironic praise scale in the *irony non-alert* group and the *irony alert* group, respectively (values for the *irony alert* group in brackets).

Achievement Measurement System 2 (*LPS-2* [*Leistungsprüfsystem 2*]*;* Kreuzpointner, Lukesch, & Horn, 2013). The LPS-2 is a performance test for the assessment of general mental ability. It employs eleven subtests that are allocated to four of the eight dimensions proposed by Carroll's (1993) model of intelligence, namely "crystallized intelligence" (e.g., solving anagrams), "fluid intelligence" (e.g., reasoning), "visual perception" (i.e., the ability to generate and process mental representations of spatial objects, to visualize, and to detect spatial patterns, e.g., mental rotation), and "cognitive speed" (e.g., arithmetic). A general IQ score is derived by aggregating the four subscales. Internal consistencies for subtests and the four dimensions are satisfactory in the norm sample with Cronbach's alpha ranging from .72 to .95. The total internal consistency for form A (form B) is high in the norm sample, $\alpha = .96$ ($\alpha =. 97$). Split-half reliability of subtests ranges from sufficient ($r_{tt} = .81$) to high ($r_{tt} = .93$). Validity is confirmed in terms of concurrence with a range of other tests of mental ability. Furthermore, the targeted dimensional structure of the test is confirmed. The LPS-2 can be administered in groups and takes around 60 minutes to complete.

Irony/lie discrimination task (Winner et al., 1998). This task measures the capacity to attribute second-order mental state and the ability to distinguish between ironic statements

and lies. Subjects are required to read 15 short stories and to identify whether the final assertion is a lie or an ironic joke. There are eight stories involving a lie and seven stories implicating irony (in terms of intentionally and overtly uttering a counterfactual statement to a person known to be aware of the true circumstances). According to the characterization given by Winner et al. (1998), each story describes a context in which one person witnesses another individual breaking a rule sneakily (e.g., stealing food). The main difference between the two story types is that in the lie stories, the protagonist does not know that he or she had been seen doing the "sneaky action" and utters a lie to the witness to avoid getting caught. In the ironic stories, the protagonist knows he or she has been seen during the transgression and thereupon utters an ironic comment (i.e., a joke) to conceal his or her shame of being caught. For each story type (i.e., "joke" stories and lie stories), a separate score is generated by summing up participants' individual false negative decisions (i.e., the discrimination errors).

State-Trait Cheerfulness Inventory (*STCI*; Ruch et al., 1996). The STCI is a questionnaire measure for the components of exhilaratability as the temperamental basis of the sense of humor. The trait version (STCI-T) encompasses three scales assessing *cheerfulness* (e.g., "I have a "sunny" nature."), *seriousness* (e.g., "I prefer people who communicate with deliberation and objectivity.") and *bad mood* (e.g., "Even if there is no reason, I often feel ill-humored."). In current study a 60-item short form of the STCI-T was used. The questionnaire assesses the endorsements of statements on a four-point scale (ranging from 1 = "strongly disagree" to 4 = "strongly agree"). Internal consistencies in the present sample were comparable to the ones in the construction sample reported by Ruch et al. (1996) with Cronbach's alpha ranging from .80 (seriousness) to .95 (bad mood).

Statements of *Benevolent and Corrective Humor* (*BenCor*; Ruch & Heintz, 2016a). The BenCor is a list of statements assessing two virtue-related facets of the sense of humor. Six statements are used for benevolent humor (e.g., "Even when facing unpleasant events I

can keep my distance and discover something amusing or funny in it") and corrective humor (e.g., "I caricature my fellow humans' wrongdoings in a funny way to gently urge them to change"), each. They were answered on a 7-point Likert scale ranging from 1 ("strongly disagree") to 7 ("strongly agree"). Internal consistencies in the present sample were sufficient: Cronbach's alpha was .75 for benevolent humor and .78 for corrective humor.

Inventory of Minimal Redundant Scales (*MRS-25* [Inventar Minimal Redundanter Skalen], Ostendorf, 1990; 25-item short form developed by Schallberger & Venetz, 1999). The MRS-25 is a list of 25 bipolar adjectives pairs for the assessment of the Big Five personality dimensions *extraversion* (e.g., impulsive vs. restrained), *agreeableness* (e.g., affirmative vs. oppositional), *conscientiousness* (e.g., diligent vs. lazy), *emotional stability* (e.g., robust vs. vulnerable), and *culture* (e.g., inventive vs. conventional). Answers are given on a six-point scale (*very—quite—rather—rather—quite—very*). Schallberger and Venetz (1999) report high internal consistencies of the scales and evidence for the validity of the MRS-25. Internal consistencies in the present sample were satisfactory with Cronbach's alpha ranging from .72 (agreeableness) to .86 (conscientiousness and emotional stability).

Procedure

Participants were tested in two consecutive sessions. In Session 1, groups up to 30 persons completed the LPS-2 as the first part of a larger assessment battery also including measures that were unrelated to the present study in the laboratory, quasi-randomly assigned to form A or Form B, depending on their seating position (as to avoid influence by neighboring participants). Due to time constraints, all other measures were included in an online survey. Participants were assigned an individual code and provided with an invitation containing an URL directing them to the online survey (Session 2). Within seven days after Session 1, participants logged in and indicated their personal code for matching purposes. In Session 2, participants first completed the TOVIDA quasi-randomly assigned to one of two

conditions: Half of the groups tested in Session 1 were given a definition of verbal irony and were instructed to watch out for irony, i.e., they were told that some of the scenarios they were about to appraise contain verbal irony whereas others do not (*irony alert* condition). The other half took the test naïve to its true intention (*irony non-alert* condition), i.e. there was no mention of the possible occurrence of verbal irony. Subsequently, STCI-T, the Big Five measure (MRS-25), the sense of humor measure (i.e., the BenCor), and the irony/lie discrimination task by Winner et al. (1998) were completed.

Results

Is the detection of ironic criticism and ironic praise associated with abilities and traits?

The correlations between the two subscales of the TOVIDA and the other measures are given in Table 4, for the irony non-alert and the irony-alert group separately. As Table 4 shows, the ironic criticism scale was correlated substantially with the ironic praise scale but not correlated significantly with the other measures in the *irony non-alert* group. However, there was a trend for an association between the ironic criticism scale and the visual perception dimension of the LPS-2 (i.e., spatial ability), the performance in the ironic items (i.e., the joke stories) of the irony/lie discrimination task by Winner et al. (1998), and culture. In the *irony alert* group, again the ironic criticism scale was correlated substantially with the ironic praise scale. Furthermore, as expected, there was an association between the ironic criticism scale and the performance in the ironic items of the irony/lie discrimination task by Winner et al. (1998). Furthermore, there was also a trend for an association between the ironic criticism scale and emotional stability. In line with the expectations, among the self-report measures, bad mood and benevolent humor showed a significant relation to the ironic criticism scale and there was a trend for an association with cheerfulness. Furthermore, there was also a

trend for ironic criticism detection showing an association with agreeableness and emotional stability.

As expected, the ironic praise scale was significantly correlated with intelligence in terms of fluid intelligence and with the performance in the ironic items of the irony/lie discrimination task in the *irony non-alert* group. Furthermore, there was a trend for an association with visual perception and culture for the ironic praise scale. In the *irony alert* group, the ironic praise scale was associated with intelligence in terms of the LPS-2 dimension visual perception (and there was also a trend for an association with the fluid intelligence dimension). Furthermore, the ironic praise scale again was negatively correlated with the number of errors made in the ironic items of the irony/lie discrimination task by Winner et al. (1998). Among the scales of the self-report measures, emotional stability, cheerfulness, bad mood, benevolent humor, and corrective humor showed significant correlations with the ironic praise scale. Furthermore, there was also a trend for an association with extraversion for the ironic praise scale in this group.

Are there unique correlates for ironic praise beyond ironic criticism?

Next, it was tested whether in the study of irony detection correlates ironic praise generates meaningful variance that contributes a surplus value over the meaningful variance found for ironic criticism. Therefore, partial correlations were computed between the ironic praise detection scale and the external variables while controlling for individuals' ironic criticism detection scores. The partial correlations are given in Table 4. As can be seen in Table 4, in the *irony non-alert* group, ironic praise correlated positively with *fluid intelligence* and negatively with the error rate in the irony items of the irony/lie discrimination task even beyond the influence of the variance shared with ironic criticism detection. In the *irony alert* group ironic praise correlated positively with the *visual perception* dimension of the intelligence test, trait cheerfulness, trait bad mood (in a negative direction), and corrective

Table 4

Correlations Between Irony Detection Scores and the Personality and Ability Measures (Study 3).

	TOVIDA test instruction					
	Irony non-alert			Irony alert		
Personality and ability measures	IC	IP	IP_p	IC	IP	IP_p
TOVIDA IP	.52*	-	-	.46*	-	-
Intelligence (LPS-2)						
Crystallized intelligence	.03	.04	.03	.00	.10	.11
Fluid intelligence	.05	.25*	.26*	-.04	.17	.21
Visual perception	.15	.16	.09	.07	.28*	.28*
Cognitive speed	-.07	.09	.14	.00	.14	.16
General IQ	.06	.22*	.22*	.03	.23*	.24*
Irony/lie discrimination						
Irony (joke stories)	-.17	-.27*	-.23*	-.24*	-.30*	-.22
Non-irony (lie stories)	.09	.01	-.04	.10	.02	-.03

Table 4 continues.

Table 4 continued.

Personality and ability measures	TOVIDA test instruction					
	Irony non-alert			Irony alert		
	IC	IP	IP_p	IC	IP	IP_p
Big Five						
Agreeableness	.11	.04	-.02	.17	.02	-.06
Conscientiousness	.05	.02	-.01	.05	.12	.12
Emotional Stability	.04	.03	-.01	.17	.24*	.18
Extraversion	.06	.05	-.03	.01	.17	.20
Culture	.14	.19	.14	.04	.01	.03
Temperamental traits						
Cheerfulness	-.03	-.08	-.07	.20	.29*	.23*
Seriousness	-.08	-.04	-.01	-.07	-.04	-.01
Bad Mood	-.03	.01	.04	-.34*	-.35*	-.23*
Sense of Humor						
Benevolent Humor	.11	.07	.03	.24*	.30*	.22
Corrective Humor	-.05	-.06	-.06	.09	.26*	.25*

Note. n = 97-103 irony non-alert individuals. n = 80 irony alert individuals. IC = ironic criticism scale of the TOVIDA, IP = ironic praise scale of the TOVIDA. Sense of Humor = scales of the BenCor. IP_p= partial correlations with IP controlling for the influence of IC.

* $p < .05$ (two-tailed).

humor over and above the variance that the ironic criticism scale shared with ironic praise and these variables.

Discussion

The findings of Study 3 indicate that assessing the detection of ironic praise can provide a surplus value over the detection of ironic criticism. Ironic praise detection can be seen as more challenging than the detection of ironic criticism in terms of numerically higher associations as well as significant partial correlations with the intelligence measure when the influence of the aptitude to detect ironic criticism was controlled for.[14] Hence, ironic praise detection appears to be dependent on mental ability to a certain degree, which is in line with previously reported findings on the role of intelligence in irony detection (e.g., Mitchley et al., 1998). However, considering the numerical size of the correlations, ironic praise detection aptitude can be seen as distinct from intelligence. Furthermore, as expected, it was found that the detection of ironic praise was uniquely associated with corrective humor, while ironic criticism was related only to benevolent humor. Also, cheerfulness played a unique role in the detection of ironic praise. Possibly increasing the readiness to process humorous meta-

[14] Differential associations between the two scales of the TOVIDA and the intelligence variables could be explained by differences in average item difficulty. As ironic praise items were more frequently appraised as ironic than ironic criticism items in Study 2, it is possible that the lack of association between the ironic criticism scale and intelligence hence might be an artifact created by higher ambiguity of the materials. This may be the case because intelligence may foster irony detection only when items have a low ambiguity.

messages or playful cues in ironic teasing, a cheerful temperament hence can be assumed to facilitate the detection of irony, foremost in the form of ironic praise.

The Big Five personality traits were largely unrelated to irony detection scores except for a correlation between the ironic praise scale and *emotional stability*. It can be assumed that emotionally stable individuals have a higher readiness to reject the uttered criticism in what is literally said and recognize the more benevolent nature of what is ironically implied in the ironic praise items, compared to individuals low in emotional stability (who in turn may not "get over" the criticism or insult uttered in ironic praise). Although there was also a trend for an association between the irony detection scores on the one hand and culture and agreeableness on the other, the Big Five can be seen as less relevant for irony detection than narrower and more humor-related traits. Moreover, participants' scores in the TOVIDA converged with their scores in the ironic items of the irony/lie discrimination task, indicating convergent validity of the TOVIDA.

Do ability and personality variables interact in irony detection?

As an exploratory analysis complementing our correlational analyses, we wish to address the possibility that ability and personality variables interact in irony detection. To illustrate, although intelligence was found as positively related to irony detection, there might be highly intelligent individuals who still perform poorly in irony detection because they lack the requisite personality traits facilitating irony detection. Guided by the findings displayed in Table 4, we explored the data from Study 3 to see whether interactions between intelligence and personality could be found to predict irony detection beyond the main effects of the separate variables. Indeed, this assumption was found to hold true in one of the cases that we studied: in the irony-alert sample the interaction between the spatial ability dimension of the LPS-2 (i.e., *visual perception*) and *benevolent humor* predicted ironic praise detection

significantly by explaining incremental variance beyond the main effects of the single predictors.

A hierarchical regression analysis with two steps was computed with the ironic praise detection score as the criterion. In Step 1, *visual perception* (β = .25) and *benevolent humor* (β = .26) were significant predictors, $F(2, 77) = 6.70$, $p = .002$. As it turned out, the interaction term (computed as the simple multiplication of *visual perception* and *benevolent humor* scores) explained a significant increment of criterion variance when added to the equation in Step 2, $F(3, 76) = 7.27$, $p < .001$; $\Delta R^2 = .075$, $p = .008$. As a possible interpretation of this finding, intelligence could be seen as a necessary but not sufficient condition for irony detection, as irony detection may be facilitated by individuals' cognitive ability only if individuals have enough sense of humor to successfully deal with irony. The inverse may also be true: the sense of humor may only manifest itself in irony detection performance if individuals have the necessary ability to successfully deal with its cognitive demands.

General Discussion

Our findings support the assumption that the detection of ironic criticism and the detection of ironic praise can be found as two intercorrelated but still discriminant facets of irony detection aptitude. Furthermore, our findings substantiate the assumption that ironic praise is useful beyond ironic criticism: applied in an investigation of ability and personality correlates, the detection of ironic praise was found to be uniquely associated with certain variables (i.e., intelligence, trait bad mood, trait cheerfulness, and the corrective facet of the measure of the sense of humor,) beyond the influence of ironic criticism detection aptitude.

Extrapolating our findings, we may propose assumptions as to *why* more intelligent individuals high in cheerfulness and low in bad mood with high scores in benevolent and corrective humor may have a higher readiness to detect the irony in ironic praise. Maybe they

are more able or ready to (a) reason and infer the meta-message of an ironic praise (i.e., *fluid intelligence*), (b) generate an easily interpreted mental "image" of the background of an ironic remark (i.e., *visual perception* as the ability to generate mental representations, to visualize, and to detect patterns), (c) take into account playful and humorous communicative intentions in terms of the processing of exhilarant stimuli (i.e., high trait cheerfulness and low trait bad mood), (d) have a smiling attitude toward the imperfections of life (e.g., human weakness) and know how to deal with them by using *benevolent humor* (i.e., in terms of the principle "it takes one to know one"), and (e) expose transgressions of morally valued social rules by using irony with satirical meta-messages in order to educate and better social others (i.e., the tendency to produce, to enjoy, and to make sense of *corrective humor*).

The role of irony alertness

There was an irregularity in the findings of Study 3 (which, however, occurred in a quite constant fashion): in the *irony non-alert* group, the association between the personality variables and irony detection was not evident compared to the *irony alert* group. As a possible explanation for this finding, participants in the *irony non-alert* sample may have been biased towards expecting a bona fide communication mode, as in the given psychological assessment situation a serious state of mind may have been induced. This consideration may have an implication for the assessment of irony detection in general terms, as in many of the pre-existing measurement procedures for the assessment of irony detection irony alertness is reduced by not mentioning to participants that the stimuli they are about to encounter contain irony and by using indirect measurement (i.e., not asking the participants directly whether they think that there is irony in a stimulus).[15] At least as far as the study of

[15] In previous studies indirect measurement of irony detection was operationalized for example by resorting to fact questions (e.g., Ackerman, 1983; Happé, 1993), using questions targeting mental states of the speaker and emotions of the target of the ironic utterance (e.g., McDonald & Pearce, 1996), asking whether it made any sense for the speaker to make the

personality and ability correlates of irony detection is concerned, it can be seen as worthwhile to further explore the benefit of maximizing irony alertness and using direct testing.

Is the TOVIDA too difficult?

In the construction of the TOVIDA we assumed that, in order to tap into the variance in irony performance among normally functioning adults, psychometrically difficult items need to be employed (as to avoid ceiling effects). Notably, there is a trade-off between item difficulty (i.e., ambiguousness of the stimuli) and test-takers' consensus as to the ironic nature of the stimuli. Certainly, the items should not be too difficult to allow for a sufficient consensus among test takers as to whether irony is present in the stimuli or not. However, a fair amount of variance (i.e., an imperfect consensus) can be argued to be admissible as this variance (a) must be expected when conceptualizing irony detection aptitude as an approximately normally distributed variable, and (b) is rooted in the nature of the construct when dealing with phenomena involving an inherent uncertainty, which—apart from irony— can also be found for example in certain knowledge domains. Accordingly, Legree (1995) for example argues in favor of a Likert-based assessment of social intelligence because of the level of uncertainty involved in the stimuli. He characterizes the challenge of assessing knowledge of ambiguous relationships when he states that "situational judgment scales attempt to simulate everyday problem situations but cannot allow the formulation of unambiguously 'correct' solutions. This ambiguity partially reflects real-world interpersonal interactions, which are often ambiguous […]" (Legree, 1995, p. 249).

The possible role of self-involvement

In the TOVIDA, test-takers have to make sense of situations containing verbal irony from an observer's perspective (i.e., with low *self-involvement*). It would also be thinkable to test irony detection performance using self-involving situations, such as when instructing test

target utterance (e.g., Langdon & Coltheart, 2004), or rewording the use of irony as "joking" (e.g., Winner et al., 1998).

takers to place themselves into the respective situation as if they would encounter them in real life. Importantly, this may lead to certain variables coming into play more prominently as correlates of irony detection performance. For example, self-involvement may accentuate the association between ironic praise detection and *emotional stability*. If a specific instance of ironic praise is an interpersonal evaluation, emotionally unstable individuals may be more attached to the negative interpersonal valence of the verbatim utterance (which can occur in the form of a mock critical offense) and hence may be less prone to reject the literal interpretation of the ironic remark—and importantly so this mechanism may be accentuated as self-involvement in the assessment of irony detection increases. This consideration may also apply to certain other traits, such as self-esteem or the fear of being laughed at (i.e., gelotophobia; cf. Ruch, Hofmann, Platt, & Proyer, 2014). For example, because of their general belief to be inherently ridiculous and deficient, gelotophobes may be sensitive to derisive ironic criticism especially when self-involvement is high. Accordingly, future studies investigating traits relevant to derisive criticism or offense in irony detection should explore the benefit of self-involving test stimuli and instructions.

Conclusions

Ironic criticism and ironic praise can be seen as separate scales in irony detection. The two types of irony were differently related to ability and personality variables, as ironic praise detection showed unique associations with intelligence and certain traits. Hence,—at least as far as the stimuli used in our investigation are concerned—ironic praise can be postulated to generate variance with surplus meaning beyond the variance generated by ironic criticism in irony detection. Consequently, ironic praise as the less "prototypical" and formerly neglected type of irony and can be postulated as especially important to include when studying the role of ability, personality, and humor in irony detection.

PART III

The role of humor-related traits and broad personality dimensions in irony use

Richard Bruntsch and Willibald Ruch

The official citation that should be used in referencing this material is:

Bruntsch, R., & Ruch, W. (2017). The role of humor-related traits and broad personality dimensions in irony use. *Personality and Individual Differences*, *112*, 139–143. doi:10.1016/j.paid.2017.03.004

Abstract

As humor has pervasively been postulated as a function of irony, humor-related traits such as the joy of laughing at others (i.e., katagelasticism) or trait seriousness can be assumed to predict who is more and who is less inclined to use verbal irony—even beyond the possible effect of broad personality dimensions. For the present study, $N = 153$ subjects made responses in two different irony use measures and completed personality questionnaires. As expected, irony use scores were higher among individuals who tend to break with social conventions, joyfully expose others' transgressions, or aggressively use ridicule (i.e., individuals scoring high in psychoticism, katagelasticism, or the aggressive humor style). Moreover, irony use was more prevalent among playful individuals who tend to entertain others by joyfully exposing themselves as the butt of jokes or engaging in as-if behaviors (i.e., low-serious individuals, scoring high in gelotophilia or the histrionic self-presentation style). Using a hierarchical regression analysis, it was found that over and above redundancy katagelasticism and—unexpectedly so—the self-defeating humor style predicted irony use beyond the influence of psychoticism. Accordingly, irony may also be seen a way to hide negative feelings behind humor and to avoid dealing constructively with problems.

Keywords: Humor, Irony, Katagelasticism, Psychoticism, Ridicule, Self-defeating humor style, Seriousness, STCI

Introduction

When we use verbal irony, we typically utter something different from what we actually want to say, for example by using counterfactual utterances or stating the opposite of what we mean (e.g., Haverkate, 1990). Characteristically, we expect the addressee to see through this dissimulation and detect what we actually want to express nevertheless (cf. Groeben & Scheele, 2003). Initial studies have investigated cultural or situational factors predicting irony use—for example by comparing the prevalence of irony use between groups from collectivist and individualist cultures (e.g., Rockwell & Theriot, 2001) or between different experimental conditions (such as computer-mediated vs. face-to-face communication, i.e., Hancock, 2004). However, the investigation of irony use as an enduring tendency—stable across different settings or situational contexts—and its relation to personality traits has not been targeted yet.

Introducing an individual differences perspective in irony research, Bruntsch, Hofmann, and Ruch (2016) advocate that the tendency to use irony can be expected to systematically differ between individuals, depending on personality traits. Previous findings indicate that indeed (a) there is interindividual variance in irony use in terms of considerable standard deviations (e.g., Matthews, Hancock, & Dunham, 2006), and (b) this variance is meaningful (rather than negligible "noise") as it can be explained by trait variables relating to the utility of irony. That is, Averbeck and Hample (2008) found that individual differences variables related to interpersonal aggression, such as verbal aggressiveness, were associated with irony use.[16]

The role of aggression as a function of irony—foremost in its sarcastic form—is well established in irony research (e.g., Dews, Kaplan, & Winner, 1995). Likewise, the notion that

[16] In the course of evaluating the German adaptation of the irony use measure introduced by Averbeck and Hample (2008), we found that the irony use scale of this instrument (consisting of ten items) had a sufficient internal consistency to support the assumption that irony use varies (a) systematically between individuals and (b) to a relatively small extent within individuals in terms of an enduring tendency to use irony ($\alpha = .83$, $N = 97$).

humor is a function of irony is pervasive in the existing literature. For example, laypersons were found to view humor as a reason why irony is used (e.g., Roberts & Kreuz, 1994). Furthermore, indicating that humor is a motive for irony use, Matthews et al. (2006) found that humorous ironic options were chosen more frequently than less humorous ironic options in their irony use measure. From a theoretical stance, if humor is considered also in the disparaging part of its spectrum (cf. Zillmann, 1983), irony can be seen as related to humor because it is suitable for victimizing others while having a humorous effect on bystanders (cf. Garmendia, 2014).

Accordingly, Bruntsch et al. (2016) hypothesized that a range of humor-related personality traits may be associated with the use of irony. Namely, these are *trait seriousness* and *trait bad mood* (Ruch, Köhler, & van Thriel, 1996), certain dispositions to laughter and ridicule (i.e., *katagelasticism* and *gelotophilia*; Ruch & Proyer, 2009), *the histrionic self-presentation style* (i.e., the inclination to engage in *as-if* behaviors; Renner, Enz, Friedel, Merzbacher, & Laux, 2008), and *the sense of humor* (in terms of stable interindividual differences in the way people react to and produce humor and a cheerfully composed attitude toward life; see Ruch, 1998).

As an approach to conceptualize the sense of humor, Martin, Puhlik-Doris, Larsen, Gray, and Weir (2003) introduced four humor "styles" relating to individual differences humor use: the affiliative, self-enhancing, aggressive, and self-defeating humor style. As irony was reported relate to aggression and humor, the *aggressive humor style* presumably is relevant for the present research question. Furthermore, as irony can be seen as a means of interpersonal bonding when used in ironic teasing (cf. Keltner, Capps, Kring, Young, & Heerey, 2001), the *affiliative humor style* (as the tendency to use humor to enhance one's relationship with others in a benign and self-accepting way; Martin et al., 2003) is expected

to be positively associated with irony use. We did not have any hypotheses for the other two (i.e., *self-enhancing*, *self-defeating*) humor styles.

Eysenck's personality dimensions may be used to control for the influence of broad personality dimensions (i.e., extraversion, neuroticism, and psychoticism; cf. Eysenck & Eysenck, 1991). They were preferred to the Big Five dimensions of personality in the present study because psychoticism is not entailed in the Big Five model but may have a special importance for our research question. A person scoring high in psychoticism is characterized as aggressive, cold, egocentric, impersonal, impulsive, antisocial, unempathic, creative, and tough-minded (cf. Eysenck, 1992). Hence, individuals with higher (vs. lower) scores in psychoticism may be less (vs. more) inhibited by social norms (such as kindness and sympathy) and hence be more (vs. less) prone to use irony in order to expose and ridicule others' transgressions. Extraversion can be seen as relevant for the present research question because humor behaviors (including irony use) involve positive emotions, which are typically more frequent in extraverts than in introverts (cf. Ruch & Deckers, 1993). Likewise, emotionally stable individuals may be more inclined to risk offending others by using playful provocation in terms of ironic teasing (cf. Keltner et al., 2001) than neuroticistic individuals.

The aim of this paper is to explore whether (a) broad personality dimensions and humor-related traits predict irony use and, if this is the case, whether (b) humor-related traits have an incremental value over broad personality dimensions in this prediction. As a prerequisite, it is expected that there is systematic interindividual variance in irony use in terms of an enduring tendency manifesting itself in a substantial inter-correlation between scores of different measures of irony use. It is hypothesized that traits facilitating humor-related behavior (i.e., gelotophilia, katagelasticism, the aggressive and the affiliative humor style, and the histrionic self-presentation style) as well as trait bad mood will predict irony use in a positive direction, just as traits impeding humorous behavior (i.e., trait seriousness)

will predict the use of irony in a negative direction. Furthermore, it is expected that irony use correlates positively with extraversion and psychoticism but negatively with neuroticism. Due to the ubiquitous assumption that humor and irony overlap, the relevant humor-related traits are expected to explain incremental variance in irony use behavior beyond the possible influence of Eysenck's personality dimensions.

Method

Sample.

Participants were recruited via university mailing lists and social platforms. The sample consisted of 153 German-speaking subjects (39 male [25%]; age: 18 to 69 years, $M = 26.4$, $SD = 10.4$).

Instruments

Irony use measures

Forced choice irony use measure. The forced choice measure for the assessment of irony use is taken from Matthews et al.'s (2006) materials, translated into German (using a translation and back-translation procedure), and adapted. Participants have to give a response to each of eight different situations by choosing between four (i.e., one ironic and three non-ironic) options. There are two response options added to the original method by Matthews et al. (2006, Experiment 4), which serve as distractors: one is designed as a non-ironic aggressive response and one designed as a non-ironic humorous response. With only two options provided, it can be seen as hard to distinguish whether participants choose a response because they are appealed by this response or because they reject the other response. There are four situations providing a mock positive evaluation of negative circumstances (ironic criticisms), and four situations that provide a mock negative evaluation of positive circumstances (ironic praise) as the ironic response option. Participants have to indicate what they most likely

would say in place of the respective person in the given situation. Scenarios were also adapted to be eliciting the use of irony by including "antecedents" (cf. Kreuz & Glucksberg, 1989) in the scenario descriptions that hinted at violations of positive expectancies or norms that give reason for conveying a critical attitude via an ironic remark (cf. Garmendia, 2014). It may be illustrating to take a look at the following scenario used in the present study: "*You and Chris have been friends all through university. As long as you've known him, Chris has always been very careful about his appearance and often wears designer clothes. At your graduation ceremony, no one was surprised when Chris showed up wearing a new Armani suit under his gown. During the ceremony he kept fumbling around with his tie knot, discontented, to adjust it to perfect fit."*. The last sentence was added to Matthews et al.'s (2006) original scenario in order to make the ironic response option occur more characteristically (as in this case vanity can be seen as a transgression of a social norm that is suitable to be addressed by ironic praise in a teasing manner). The response options provided for this scenario read as follows: "*Gosh Chris, you're looking a little scruffy for the big ceremony.*" (ironic praise), "*That's really a great suit, Chris.*" (non-ironic praise), "*Gosh Chris, don't be such a peacock!*" (non-ironic aggressive criticism), and "*I wish I was able to tie a tie knot like this in the first place!*" (non-ironic humorous response). A total score counting participants' choices of ironic responses (with a minimum possible value of zero and a maximum possible value of eight) was computed.

Rating-based irony use measure. The rating-based measure for the assessment of irony use is taken from Averbeck and Hample (2008) and translated into German (using a translation and back-translation procedure). Participants are provided with a detailed scenario in which they are asked to assume the perspective of a person who shares a flat with a friend who—among other misconducts—is messy and does not pay the bills for the flat. Participants are asked to indicate on a Likert-type scale ranging from 1 ("not likely at all") to 4 ("very

likely") how likely they were to use each of 20 utterances. Among the utterances each of ten criticisms (e.g., regarding noise in the flat or unpaid bills) is communicated once ironically and once non-ironically. In the total score, the responses to the ten ironic items were corrected for the general tendency to utter criticisms by subtracting the average endorsement of the ten literal criticisms (α = .79) from the average endorsement of the ten ironic criticisms (α = .81).

Personality and trait measures

The *Eysenck Personality Questionnaire-Revised* (*EPQ-RK*) in the German adaptation and short form by Ruch (1999) of the English version (*EPQ-R*; Eysenck & Eysenck, 1991) was used for the assessment of psychoticism, extraversion, and neuroticism in terms of Eysenck's model of personality. The 50-item questionnaire contains the three named content scales and a lie-scale. The answer format is dichotomous ("*yes"* or *"no"*). Internal consistencies in the present sample ranged from α = .52 for psychoticism to α = .85 for neuroticism.

State-Trait Cheerfulness Inventory (*STCI-T<60>*; Ruch et al., 1996). The STCI-T<60> is the trait version of the established questionnaire for the assessment of the components of exhilaratability as the temperamental basis of the sense of humor. The STCI-T<60> encompasses three scales assessing the traits *cheerfulness*, *seriousness*, and *bad mood* with 20 items each. Items are answered on a four-point scale (ranging from 1 = "strongly disagree" to 4 = "strongly agree"). Internal consistencies in the present sample ranged from α = .81 (seriousness) to α = .95 (bad mood).

Humor Styles Questionnaire (HSQ; Martin et al., 2003; German version by Ruch & Heintz, 2016b). The HSQ consists of 32 items measuring four humor "styles" (the *affiliative*, *self-enhancing*, *aggressive*, and *self-defeating* humor style). The instrument uses a seven-point Likert-scale from 1 ("totally disagree") to 7 ("totally agree"). Internal consistencies in the present sample ranged from α = .75 (*self-defeating*) to α = .81 (*self-enhancing*).

The *PhoPhiKat-45* (Ruch & Proyer, 2009) is a questionnaire for the assessment of gelotophobia (i.e., fear of being laughed at), gelotophilia (i.e., joy of being laughed at), and katagelasticism (i.e., joy of laughing at others; 15 items each) employing a four-point answer format (1 = "strongly disagree" to 4 = "strongly agree"). Internal consistencies in the present sample ranged from α = .85 (katagelasticism) to α = .88 (gelotophobia).

The *As-If Scale* (*AIS*, Renner et al., 2008) assesses the histrionic self-presentation style as an inclination to humorously shape everyday interactions by using as-if behaviors such as acting, role-playing, or imitating. It consists of eight items and employs a four-point answer format (1 = "does not apply at all" to 4 = "totally applies"). The internal consistency of the AIS in the present sample was α = .81.

Procedure

Participants were tested via an online survey. As the completion of the irony use measures was reasoned to be more susceptible to priming and carry-over effects (possibly generating artificial co-variance between the irony use measures and the trait measures) than the assessment of stable and enduring traits, irony use was assessed prior to personality traits. Participants first completed the forced choice and then the rating-based irony use measure. Subsequently, the EPQ-RK, the STCI-T<60>, the HSQ, the PhoPhiKat-45, and the AIS were completed. Participants were instructed to fill in the irony use measures by responding as if they would encounter the situations in real life. It was not mentioned to participants that the survey would include ironic stimuli.

Analyses

To arrive at an aggregated measure of irony use, the scores of the forced choice irony use measure (M = 1.32, SD = 1.17) and the rating-based irony use measure (M = -1.68, SD =

0.78) were standardized by *z*-transformation and averaged in one total score. A high level of aggregation was preferred to analyzing single scores, as (a) we did not have hypotheses concerning differences between the two irony use measures as to their associations with the personality traits, and (b) we desired to reduce possible measure-specific effects.

In order to explore personality correlates of irony use, partial correlations between the irony use score and the individual differences measures were computed by statistically controlling for the influence of gender as a possible confounding variable. Furthermore, to simultaneously test (a) which personality traits explain the tendency to use irony in the given irony use measures as unique predictors, and (b) whether the predictive power of humor-related traits goes beyond the influence of broad personality dimensions, a hierarchical regression analysis was computed with the irony use score as the criterion and the trait scales as predictors. To account for its possible influence, participants' gender was entered as a predictor before all other predictors on the first hierarchical level in Step 1. On the second hierarchical level, the scales of the EPQ-RK were entered in Step 2 of the analysis (method: stepwise, i.e., in an iterative procedure the strongest predictor is taken into the equation first and it is tested whether an additional variable explains incremental variance in the criterion beyond what is already accounted for by the previously added predictor[s]). On the third hierarchical level (Step 3), participants' scores in the STCI-T<60>, the HSQ, the PhoPhiKat-45, and the AIS were given into the analysis in addition to the predictors added to the equation in Steps 1 and 2 (method: stepwise).

Results

The correlation between the two irony use measures was $r(151) = .44$, ($p < .001$, 95% CI = [.30, .56]). The bivariate correlations between the irony use score and the trait measures are given in Table 5.

Table 5

Descriptive Statistics for Questionnaire Scores and Correlations Between the Irony Use Score and Personality Measures.

	Questionnaires		Irony use score	
	M	*SD*	$r_{partial}$	95% CI
Eysenck Personality Questionnaire				
Psychoticism	0.23	0.14	.33***	[.16, .48]
Extraversion	0.67	0.26	.14	[-.03, .29]
Neuroticism	0.39	0.28	-.04	[-.18, .11]
Lie Scale	0.23	0.17	-.10	[-.24, .05]
Trait Cheerfulness Inventory				
Cheerfulness	3.15	0.48	.00	[-.19, .19]
Seriousness	2.48	0.38	-.25**	[-.40, -.09]
Bad Mood	1.91	0.60	.09	[-.09, .26]
Humor Styles Questionnaire				
Affiliative	5.56	0.91	-.04	[-.24, .15]
Self-enhancing	4.54	1.05	.09	[-.09, .26]

Table 5 continues.

Table 5 continued.

	Questionnaires		Irony use score	
	M	*SD*	$r_{partial}$	95% CI
Aggressive	3.54	1.07	.32***	[.17, .46]
Self-defeating	3.45	1.02	.31***	[.16, .45]
Dispositions to laughter and ridicule				
Gelotophobia	1.98	0.54	.06	[-.12, .23]
Gelotophilia	2.39	0.53	.27**	[.11, .43]
Katagelasticism	1.99	0.47	.49***	[.34, .62]
Histrionic self-presentation style				
As-if scale	2.15	0.55	.23**	[.08, .37]

Note. $N = 153$. $r_{partial}$ = partial correlations controlling for the influence of gender. 95% CI = lower and upper limit of the 95% confidence interval.

** $p < .01$, two-tailed.

*** $p < .001$, two-tailed.

As Table 5 shows, among the scales of the EPQ-RK, psychoticism correlated substantially with irony use in a positive direction whereas extraversion, neuroticism and the lie scale were not associated with irony use significantly. Among the temperamental traits measured by the STCI-T<60>, as expected, trait seriousness correlated with the irony use score in a negative

direction. Contrary to expectations, trait cheerfulness and trait bad mood were not associated with participants' irony use significantly. Among the scales of the HSQ, the aggressive humor style correlated substantially with irony use in the expected (positive) direction. Contrary to expectations, the affiliative humor style was not correlated to the irony use score. However, unexpectedly, the *self-defeating* humor style was correlated substantially to irony use (in a positive direction). Regarding the dispositions to laughter and ridicule, as expected, the irony use score was correlated positively to katagelasticism (the joy of laughing at others) and gelotophilia (i.e., the joy of being laughed at). Furthermore, as expected, the As-If Scale as the measure of the histrionic self-presentation style was positively correlated with irony use.

The results of the hierarchical regression analysis predicting the irony use score are given in Table 6. As Table 6 shows, it was found that gender was a significant predictor, with men tending to have higher irony use scores than women, in Step 1, $F(1, 151) = 14.76$, $p < .001$. Among Eysenck's personality dimensions, psychoticism was a unique predictor for irony use in the model resulting in Step 2 beyond the influence of gender, $F(2, 150) = 17.14$, $p < .001$. As it turned out, the scores of the humor-related traits given into the analysis in addition to Eysenck's personality dimensions in Step 3 resulted in two further models explaining incremental variance beyond the models in the respective previous steps. Firstly, katagelasticism entered the equation as a unique predictor in Model 3 , $F(3, 149) = 25.89$, $p < .001$. Secondly, the self-defeating humor style emerged as a predictor explaining unique criterion variance in Model 4, $F(4, 148) = 21.39$, $p < .001$. In the final model, psychoticism, katagelasticism, and the self-defeating humor style were significant predictors and together explained 35% of the variance in the irony use score (adjusted $R^2 = .35$). No other trait variable explained incremental variance in the criterion beyond these three predictors significantly.

Table 6

Hierarchical Regression Analysis Predicting the Irony Use Score.

	B	*SE*	β	ΔR^2
Step 1: Gender				
Model 1				.08***
Gender	-0.58	0.15	-.30***	
Step 2: Eysenck's personality dimensions				
Model 2				.09***
Gender	-0.47	0.15	-.24**	
Psychoticism	1.93	0.46	.32***	
Step 3: Humor-related traits				.
Model 3				.16***
Gender	-0.20	0.14	-.10	
Psychoticism	1.25	0.43	.20**	
Katagelasticism	0.79	0.13	.44***	
Model 4				.02*
Gender	-0.26	0.14	-.13	

Table 6 continues.

Table 6 continued.

	B	*SE*	β	ΔR^2
Psychoticism	1.28	0.42	.21**	
Katagelasticism	0.69	0.14	.38***	
Self-defeating humor style	0.14	0.06	.16*	

Note. $N = 153$. Gender: 1 = male, 2 = female. Total adjusted $R^2 = .35$. Humor-related traits = scales of the STCI-T<60>, HSQ, PhoPhiKat-45, and the AIS. Method: stepwise.

* $p < .05$, two-tailed.

** $p < .01$, two-tailed.

*** $p < .001$, two-tailed.

Discussion

The present study found that humor-related traits explained incremental interindividual variance in irony use beyond psychoticism as a broad personality dimension. The correlation between the two irony use measures substantiates the assumption that there is systematic interindividual variance in irony use in terms of an enduring tendency to use irony. Involving the tendency to break with social conventions (such as the norm of kindness and sympathy), psychoticism was expected to go along with a low inhibition to offend others by callously exposing their weaknesses by ridiculing irony. This hypothesis was supported by our results. Trait seriousness is defined to go along with the "preference for a sober, object-oriented communication style, meaning to say exactly what one means without exaggeration ironic/sarcastic undertones" (Ruch et al., 1996, p. 308). This notion can be seen as

supportable in the light of our findings. Furthermore, as expected, the joy of laughing at others (i.e., katagelasticism) turned out as a correlate of irony use. Katagelasticism involves the tendency to expose others to laughter, which—as irony oftentimes is responded to with laughter (Kotthoff, 2003)—may be achieved by using irony. Likewise, we found gelotophilia (the joy of being laughed at) to be correlated with irony use. This finding is in line with our expectations as individuals scoring high in gelotophilia have a tendency to joyfully expose themselves as the butt of jokes (cf. Ruch & Proyer, 2009), which can be attained by using self-directed irony. The histrionic self-presentation style involves the tendency to draw attention to the own person and entertain others by engaging in as-if behaviors (e.g., acting, imitating, or role-playing; Renner et al., 2008). In line with Renner et al. (2008), who listed irony among the forms of communication that can be examples for histrionic as-if behaviors, the AIS correlated with irony use. Furthermore, as expected, the aggressive humor style was found to be associated with irony use. This finding is in line with Martin et al. (2003), who defined a link between the aggressive humor style as the inclination to "hostile uses of humor, in which the self is enhanced by denigrating, disparaging, excessively teasing, or ridiculing others" (p. 52) and the use of sarcasm.

However, not all of the expectations were met. Other than expected, extraversion and neuroticism were not associated with irony use. Furthermore, our results do not support the assumption that trait bad mood "might also be a disposition facilitating certain forms of humor, such as mockery, irony, cynicism, and sarcasm" (Ruch & Hofmann, 2012,) p. 86). Also, contrary to our expectations, the affiliative humor style was not correlated to irony use. This is somewhat puzzling as the affiliative humor style is defined to go along with the tendency "to engage in spontaneous witty banter" (Martin et al., 2003, p. 53). As a possible explanation, the stimuli used in this study may not be accounting for irony in the form of witty banter sufficiently.

To summarize, how can typical irony users be characterized? We may suggest that irony is used both (1) in an interpersonally "cold" fashion by tough-minded and egocentric individuals (i.e., scoring high in psychoticism) who tend to break with social conventions (e.g., kindness and sympathy) and who do not to bother about the risk of "losing friends" when using humor aggressively but rather enjoy raising a laugh when exposing others' transgressions by ridicule (i.e., scoring high in the aggressive humor style and katagelasticism), and (2) in an interpersonally "warm" fashion among individuals who are prone to not taking themselves and others too seriously and to entertain themselves and others by exposing themselves as the butt of laughter, for example when humorously acting and imitating (i.e., low-serious individuals, scoring high in gelotophilia and the histrionic self-presentation style).

When redundancy was eliminated in the prediction of irony use, a less diverse picture presented itself. Among Eysenck's personality traits psychoticism was the only unique predictor and beyond its influence only katagelasticism and—unexpectedly so—the self-defeating humor style were explaining the irony use score over and above redundancy. Hence, irony may foremost be used to callously expose others' transgressions, especially by those who enjoy to laugh or to raise a laugh at someone else's expense. As an unexpected aspect, irony use may also be a way by which the self-defeating humor style is expressed. Jocular (i.e., ironic) versions of necessary criticisms on serious topics may be a way of using humor to the detriment of the self by "hiding one's underlying negative feelings, or avoiding dealing constructively with problems" (Martin et al., 2003, p. 54). For example, ironically joking about inacceptable misbehavior and unpaid bills in a shared flat may be a self-defeating way to mute the expression of one's justified anger and concern about what is a serious threat to the social integrity and the economic well-being in the scenario given in the rating-based irony use measure.

Limitations and future studies

It is noteworthy that the selection of relevant personality traits is not exhaustive in the present study and could be extended in future studies. For example, as they denote "behavior tendencies toward self-promotion, emotional coldness, duplicity, and aggressiveness" (Paulhus & Williams, 2002, p. 557), the socially aversive traits delineated by the dark triad of personality (i.e., Machiavellianism, subclinical narcissism, and subclinical psychopathy) may predict irony use in its aggressive or "cold" spectrum. Substantiating this assumption, previous studies indicate that sub-clinical psychopathy and Machiavellianism relate to the aggressive humor style (e.g., Veselka, Schermer, Martin, & Vernon, 2010), which may manifest itself in irony use as a way to enhance the self by "denigrating, disparaging, excessively teasing, or ridiculing others" (Martin et al., 2003, p. 52). Furthermore, although in the present study a high level of aggregation was preferred, it may be worthwhile to investigate the role of personality in different forms of irony in future studies, such as ironic praise and ironic criticism.

Conclusion

Our findings can help to get a picture of *who* uses irony but also give a hint on *why* irony is used. Irony use can be predicted foremost by such traits that are relevant to humor—especially when humor is considered not only in its benevolent "fun part" but also in the disparaging part of its spectrum.

General Discussion

The three studies investigated whether (a) there is interindividual variance in irony detection performance and irony use, (b) this variance is systematic (rather than negligible "noise"), i.e., whether an irony detection aptitude and an enduring tendency to use irony can be found, (c) this variance is meaningful, i.e., whether it can be explained by ability and trait variables, (d) humor can be seen as a central aspect when generating hypotheses as to which traits are most important, (e) two different basic forms of irony lead to different results when employed in a test of irony detection aptitude as to their association with humor-related traits and intelligence, and (f) humor-relevant traits explain incremental variance in irony use behavior beyond broad personality dimensions.

Overview of the Main Results and Conclusions

PART I

In PART I it was argued that irony behaviors can be considered a phenomenon native to the realm of individual differences. Based on theoretical considerations and previous empirical findings, it was shown that (a) there is interindividual variance in irony behavior, and (b) this variance can be easily related to individual differences variables such as the sense of humor, dispositions towards laughter and ridicule (e.g., gelotophobia), and general mental ability. Furthermore, it was hypothesized that there is *sense of irony* as an individual differences variable denoting maximal and typical irony behavior. The *sense of irony* is assumed to go along with inclinations towards specific affective and cognitive processing patterns when dealing with verbal irony.

PART II

In PART II, the results indicate that (a) there are two facets of irony detection aptitude, namely the aptitude to detect ironic criticism vs. the aptitude to detect ironic praise, and (b) these facets need to be scored separately in order to make more accurate predictions and conclusions when exploring personality and ability correlates of irony detection aptitude. Certain ability and personality correlates for ironic praise detection (i.e., intelligence, cheerfulness, and corrective humor) were found to have a unique association with ironic praise detection beyond the influence of ironic criticism detection aptitude, while others (i.e., trait bad mood and benevolent humor) were jointly related to both scales (i.e., ironic criticism and ironic praise) of the irony detection measure. As a possible explanation for the unique role of intelligence in the detection of ironic praise, ironic praise detection can be assumed to require a more complex mental representation of the background of the ironic remark and a more effortful cognitive search for the antecedent event (as compared to the detection of ironic criticism). Concerning the finding that cheerfulness was associated with the detection of ironic praise beyond the influence of the variance it shares with ironic criticism detection aptitude, it may be argued that trait cheerfulness has a relevance especially to the detection of ironic praise: a cheerful and light-hearted disposition may facilitate the processing of cues signaling benevolent ironic teasing and hence may help to reject the uttered negative evaluation of circumstances (which can also take the form of a mock insult) and construe it as a playful provocation. The unique role of corrective humor in ironic praise detection is in line with the notion that by exposing transgressions of social rules in a witty and playful way, corrective humor is conceptually more related to ironic praise than to ironic criticism, which in turn can be seen as the more serious and less ingenious form of irony.

PART III

The findings of PART III help to paint a picture of *who* uses irony, but they also provide a hint on *why* irony is used. Irony use can be predicted incrementally beyond brad personality dimensions by those traits that are relevant for humor—especially when humor is considered not only for its benevolent "fun part", but also for the disparaging part of its spectrum. Taken together, the findings of PART III provide support for both (1) the notion that irony can be a means of anti-social interaction when victimizing others by relational aggression, for example in the form of disparagement humor (cf. Bowes & Katz, 2011), and (2) the notion that irony can be used as a means for shaping benevolent interactions in relationships by humorous interaction, such as in the form of entertaining or playfully teasing others. However, when looking only at unique predictors, it was found that irony is used foremost in the context of (a) aggressive, cold, egocentric, impersonal, impulsive, antisocial, unempathetic, creative, and tough-minded tendencies (i.e., psychoticism), and (b) the tendency to expose others to laughter and ridicule (i.e., *katagelasticism*/the joy of laughing at others). As an additional aspect, irony use can be seen as a self-defeating way to hide one's negative emotions behind jocularity in order not to deal with conflicts more directly and more openly.

One of the findings of PART III, the relation between irony use and the histrionic self-presentation style (Renner, Enz, Friedel, Merzbacher, & Laux, 2008) can be seen as worthwhile of being discussed in somewhat more detail, as this finding substantiates pre-existing theories of irony. One of the features of irony that different theories of irony agree on is that it involves an overt pretense (cf. Groeben & Scheele, 2003). As the insincerity in ironic utterances is typically supposed to be recognized by the addressee, irony can be characterized as involving evident simulation. This assumption is supported by the finding that irony use relates to the tendency to engage in as-if behavior (i.e., histrionic self-presentation). To illustrate: In the case of echoic irony use, the ironic utterance picks

something up that was previously said—in order to reject it as ridiculous or simply not true (e.g., Sperber & Wilson, 1981). Accordingly, it could be argued that echoic irony is similar to the concept of imitation, which again is characteristic for the histrionic self-presentation style.

Strengths, Limitations, and Future Works

Strengths

The studies conducted for the present thesis contribute to the literature by introducing a novel aspect in the psychological investigation of irony. An individual differences perspective on irony was assumed and demonstrated to be a fruitful approach to investigate the phenomena of irony detection and irony use. The studies conducted for the present thesis related the tendency to engage in irony use behavior and the aptitude to successfully detect irony to trait and ability variables. Accordingly, self-report measures and performance tests were used in order to account for irony behavior in terms of a preference-related tendency, but also as an ability-related aptitude. In PART II, the aptitude to detect irony was measured by means of a newly developed performance test (i.e., the Test of Verbal Irony Detection, TOVIDA) and related to intelligence. Unlike previously used measures (such as the irony vs. lie discrimination task by Winner et al., 1998), the TOVIDA was validated as to whether (a) the design of the stimuli is valid, i.e., whether laypersons recognize that there is irony in the stimuli, and (b) the rationale of the measurement approach is valid, i.e., whether higher (vs. lower) irony detection aptitude manifests itself in higher (vs. lower) test scores. This validation was attained by using an experimental approach, i.e., by comparing groups with "artificially" increased or decreased irony detection performance. Two central aspects of irony behavior were addressed, by examining both irony detection aptitude (i.e., irony reception ability) and the tendency to irony use irony (i.e., irony production preference).

Limitations

Irony creation ability and irony appreciation

There are four domains of irony behavior that result from the combination of irony reception vs. production and irony preference vs. ability. These four domains are described in Appendix A1. These domains are irony reception ability (i.e., irony detection aptitude), irony production preference (i.e., irony use tendency), irony reception preference (i.e., irony appreciation), and irony production ability (i.e., irony creation ability). While the former two were investigated in the present thesis, the latter two were not. Just like irony detection aptitude and irony use tendency (as investigated in the present thesis), irony creation ability and irony appreciation can be expected to vary systematically between individuals. It seems worthwhile to complete the picture of which individual difference variables play a role in irony behavior by investigating irony creation ability and irony appreciation. In PART I, the suggestion was made that there is a *sense of irony* denoting individual differences in the extent to which people detect irony, produce irony, seek irony, or enjoy irony. As a central question arising from this claim that needs to be explored is whether the four domains of irony behavior are truly interrelated.

Overlearned ironic criticism

There is another aspect worth discussing in the context of the findings presented in PART II: Since ironic criticism can be seen as the more prevalent type of verbal irony, individuals' aptitude to detect ironic criticism may in part depend on the automated recognition of overlearned idiomatic phrases in conventionalized irony (cf. Schwoebel, Dews, Winner, & Srinivas, 2000; e.g., "There's no need to rush, we've got all the time in the world!", when the speaker is in a hurry) or stereotypical sayings (e.g., "Brilliant, I just love this idea!", when the speaker disapproves of something). In contrast, ironic praise as the rarer form of irony may provide a more pristine paradigm for the study of irony detection, as individuals may not

have overlearned instances of ironic praise to the same extent as instances of ironic criticism. Substantiating this assumption, Schwoebel et al. (2000) report that they observed ironic praise ("e.g., saying 'You have a hard life' to a friend going off to the Caribbean for an all-expense-paid vacation", p. 48), as less prevalent than ironic criticism in contemporary American television shows. Hence, the detection of ironic praise can be seen as more diagnostic of individuals' primary irony detection aptitude. This may be especially interesting for the investigation of correlates of irony detection, because the detection of ironic praise can be seen as less biased by automaticity and, therefore, possible associations with relevant ability and personality variables may be less distorted by experience, education, or culture than may be the case for ironic criticism. Thus, the interpretability of the results of PART II may be limited because the correlations found between the detection of ironic criticism and personality and intelligence may have been underestimated.

Measurement of the five factors of personality ("Big Five")

As an important weakness of Study 3 in PART II, the findings need to be seen as limited by the use of a relatively short measure of the "*Big Five*" personality traits (i.e., employing five items per dimension only). Although internal consistencies in the present sample were satisfactory with Cronbach's alpha ranging from .72 (agreeableness) to .86 (conscientiousness and emotional stability), it is most likely that longer questionnaires for the assessment of the five factors cover more facets of the traits. Possibly, the MRS-25 omits some of the very aspects of the traits that may be most relevant to irony detection aptitude. Hence, some of the correlational trends found in Study 3 of PART II may have been found as substantial correlations (such as the association between ironic criticism detection aptitude on the one hand and agreeableness and culture on the other), if longer and more content valid scales had been used (not to mention that reliability is higher in longer scales, reducing noise in the correlation). To conclude: the role of the five factors of personality can be seen as

worth addressing in future studies on irony detection aptitude, preferably using the original inventories rather than their short forms.

The self-defeating humor style

As regards the finding that the *self-defeating humor style* unexpectedly emerged as a unique predictor, it may be necessary to discuss possible limitations that go along with the concept of this trait. It was found earlier that the self-defeating humor style scale of the HSQ is negatively associated with self-esteem (cf. Ruch & Heintz, 2014). Furthermore, Ruch and Heintz (2014) suggest that it may be possible to disentangle the self-defeating humor style into two components: The actual humor component (i.e., in terms of making others laugh at one's own expense) and a context component (i.e., in terms of a more general self-defeating social interaction style). In a study by Ruch and Heintz (2017), this assumption was substantiated when experimentally manipulating the wording of the items of the HSQ to either contain only (or mostly) humor (as the construct-relevant content) or to change the wording by using non-humorous alternatives instead of humor (so only context content was provided): For the self-defeating scale, it was found that "the non-humorous elements determined this humor style more than humor did" (Ruch & Heintz, 2017, p. 11). It remains a question for future studies how self-esteem and irony use are related to each other, and which component of the self-defeating humor style it is that plays the more prominent role in irony use. For now, it makes sense to suggest that irony use—presumably foremost in the case of anger-deflecting ironic criticism—can be seen as a type of behavior by which the self-defeating humor style is apt to express itself.

Irony use and corrective humor

As stated in PART III, the selection of measures for humor-related traits is not exhaustive in the respective study and could be extended in future studies. For example, Ruch and Heintz (2016a) list irony among the linguistic peculiarities of corrective humor. As Ruch and Heintz

(2016a) point out, corrective humor "uses wit to ridicule vices, follies, abuses, and shortcomings with the intent of shaming individuals and groups into improvement" (p. 36). This could also be taken to suggest that irony should be seen a tool for bettering others and for educational purposes, not just for anti-social aggression or simple amusement.

Assessment beyond predefined utterances

One possible limitation of the present study is that with regard to the two irony use measures, participants had to respond to predefined ironic utterances, thereby possibly over-estimating the frequency of irony use. Although an over-estimation of the *absolute* frequency of irony use may not affect the *relative* association between irony use *tendency* and personality traits, future studies could also employ measures of irony use with an open-ended response format, which was demonstrated to be a useful approach by Dress et al. (2008).

Some thoughts on the adaptation of the irony use measures

Another issue that requires discussing is that of the adaptation of the instruments used for the assessment of irony use tendency. In PART II it was not discussed whether the newly adapted versions of the instruments were comparable with the original instruments as to their psychometric properties. Although originally, i.e., in the studies by Matthews et al. (2006) and Averbeck and Hample (2008), the instruments were not intended to be evaluated in terms of their psychometric properties (as they were used as experimental tasks or outcomes), it is possible to gather some of the information necessary for such a comparison from their research reports. As described in PART III, the distribution of the scores of the forced choice irony use measure indicated that irony was used on average only somewhat more than once by each individual across the eight items ($M = 1.32$, $SD = 1.17$). That means that, for example, most participants chose the ironic response option, while mostly they responded by choosing one of the three non-ironic options provided. It is noteworthy that Matthews et al. (2006) provided only two response options in their experiment (Experiment 4). When, in a

first experimental condition (between subjects), their scenarios were framed to make the ironic response option (e.g., "Great shot, Dave.") to appear as an ironic criticism (i.e., when the scenario involved an act with a negative valence), their participants on average preferred to ironic response in somewhat more than half of the eight scenarios ($M = 4.31$, $SD = 1.14$). When, in a second experimental condition, the scenarios were phrased to include a positive act, the ironic response options appear as an ironic praise (e.g., "Wow Dave, you suck."), on average two out of the eight items were answered with an ironic response ($M = 2.00$, $SD = 1.75$). Hence, it seems almost natural that the frequency of irony use was lower in PART III, as there were two more non-ironic response options for each item, one of which may have been especially appealing in terms of humorousness (i.e., the non-ironic humorous response). Furthermore, in PART III, ironic criticism scenarios and ironic praise scenarios were counterbalances within subjects instead of between subjects (in order to attain a content valid instrument featuring both basic forms of irony). Also, using the measure as an experimental task, Matthews et al. (2006) did not report internal consistencies. Taken together, it is hard to compare the instrument used by Matthews et al. (2006) and the one used in PART III. Taking into account that generally the items of the forced-choice measure used in PART III psychometrically were rather difficult (as indicated by the low means of the score) it is recommendable to revise either the stimuli or the response format in future studies (for example by using rankings) to avoid a floor effect. By means of rankings, it would be possible to attain a score for each option of each item, which would make the estimation of reliability by means of internal consistency possible (as for the forced-choice format, information about the preferences beyond the "first" choices are missing).

As regards the rating-based measure for the assessment of irony use tendency, again, there were substantial deviations between the original measure as used by Averbeck and Hample's (2008) and the adaptation as used in PART III. Most importantly, Averbeck and

Hample's (2008) report does not allow for a strict replication of their procedure, as they described neither (a) which instruction they used for their participants' task nor (b) the response format they employed. Most likely, telling from the descriptive statistics they report, they used either (a) the frequency of statements endorsed in a dichotomous response format, or (b) the frequency with which an ironic response was preferred over a non-ironic response, both leading to a score with a possible range between "1" and "10". Telling from the distribution of their "irony endorsement" score in the "roommate scenario" (M = 2.10, SD = 2.58), ironic utterances were endorsed less often than literal utterances (M = 8.08, SD = 2.10). This corresponds to the findings of PART III, as individuals on average endorsed literal criticism more strongly than ironic utterances in terms of the likelihood with which they would make the statement in real life, as indicated by the negative mean of the difference score (M = -1.68, SD = 0.78; i.e., individuals' scores averaging the ratings of the ten literal utterances were subtracted from individuals' scores averaging the ratings of the ten ironic utterances, leading to a possible range of "-4" to "+4", given the four-point answer format). Using the measure as an experimental task, Averbeck and Hample (2008) did not report internal consistencies, which is why this psychometric property cannot be compared between PART III and their study. The internal consistency of the rating-based measure as found in PART III at least makes it thinkable to recommend it as a reliable instrument for the use in future studies. The coefficient alpha found for the *literal* criticism scale (α = .79) was equally high than the one found for *ironic* criticism scale (α = .81), and the two scales were negatively intercorrelated (r[152] = -.31, $p < .001$; as found in a re-analysis of the data), which may be seen as an indicator for the construct validity of the measure (as the preference for literal communication is supposed to be inverse to irony use tendency). Not least, the two measures were substantially intercorrelated, indicating concurrent validity.

To conclude: the adaptation of the two irony use measures used in PART III should be seen primarily as an adaptation of instruments originally used as experimental outcomes (i.e., measures for a dependent variable) into individual differences measures. That is, the term *adaptation* should not be confused with its meaning in a narrow sense within the field of psychological assessment, which mostly denotes the procedure with which questionnaires are translated and their psychometric properties are evaluated in a different language and culture. Nonetheless, in PART III it was found that there are indicators of good reliability and validity for the two measures, making it possible to recommend the respective approaches for future studies.

Future Works

The "mixed-motive hypothesis" of irony use

While PART III proposed an *interpersonally "cold" vs. "warm"* taxonomy of irony use, the characteristics of irony may not always be that clear-cut. Therefore, we could propose that, rather than conceptualizing the effects of irony use as a dichotomy, ironic communication may run along the same continuum that was suggested for conversational joking by Boxer and Cortés-Conde (1997). As these authors suggest, "conversational joking, when it involves teasing, functions on a continuum that ranges from bonding to nipping to biting" (Boxer & Cortés-Conde, 1997, p. 276). If the social function of irony is conceptualized as a continuum, there may be ironic interactions in which interpersonally "cold" and "warm" intentions are mixed, for example, when engaging in playful provocation of close others in terms of ironic teasing. Here, teasing can be seen as a *mixed-motive* interaction aiming to both (a) convey mockery, *and* (b) signal social inclusion. This may be the case because teasers demonstrate that they think that the relationship with the communication partner is beyond the "politeness stage" when everyone is trying to be respectful and friendly in order not to offend each other

(cf. Norrick, 1994). As Kreuz, Long, and Church, 1991) point out, irony makes it possible for the speaker to fulfill communication goals that would be more difficult to fulfill literally, if not impossible. Conveying both a positive *and* a negative message at the same time in terms of multiple communication goals may be one of the core achievements of verbal irony. Or, as Norrick (1994) puts it: "sarcasm […] can express both aggression and solidarity - aggression in the message, attacking others for their foibles and errors, and solidarity in the metamessage, including others in a playful relationship with increased involvement" (p. 423).

It may be worthwhile to address the assumptions going along with this *mixed-motive hypothesis* in future studies. For example, it may be possible to explore the interaction between personality traits and experimental variations of social consequences of ironic utterances (such as ostracizing or hurting others' feelings). As one central assumption, it may be expected that traits relating socially "cold" use of irony come into play more prominently if irony use has more serious negative social consequences (compared to a condition in which the social threat to the target of irony is low). Individuals scoring high in psychoticism can be expected to have a lower social dependency and individuals scoring high in katagelasticism may be dominated by the motivation to ridicule others (rather than to make friends), compared to low scorers in the respective trait. This is why we may expect high scorers to be less influenced by variations of social damage than low scorers. Conversely, among individuals scoring low in psychoticism or katagelasticism, irony use can be expected to decrease more strongly than among high scorers as the social consequences of the irony become more severe.

Interplay between irony detection aptitude and personality variables in irony use

Extending the research question of the study conducted in PART III, it may not only be a question of personality whether or not irony is preferred over non-irony in communication, but also a question of irony detection aptitude. It may be hypothesized that the more that

individuals are able to detect irony, the more it will depend on theoretically relevant personality traits whether they prefer to use irony or not. In other words, irony detection aptitude can be hypothesized as a necessary but not sufficient condition for irony use, and there may be an interaction between irony detection aptitude and irony-relevant personality traits. More specifically, the capacity to detect irony may be a prerequisite to be able use irony; however, it may not fully explain irony use when individuals have the prerequisite ability to use irony (but due their scores in certain traits do not make use of it).

Gelotophobia and false alarms in irony detection

As suggested in PART I of the present thesis, an individual differences variable likely to affect the misinterpretation of verbal irony is the fear of being laughed at (gelotophobia). The fear of being laughed at is a personality trait characterized by the bias to experience a broad range of social interactions involving humor (which gelotophobes generally misperceive as put-down humor if directed at them) and laughter (generally perceived as victimizing) as hurtful attacks, ridicule, and contempt (see Ruch, Hofmann, Platt & Proyer, 2014, for a review). Gelotophobia has received considerable attention in a substantial body of research (see Ruch et al., 2014). In previous studies it was found that individuals with the fear of being laughed at are inclined to misperceive social situations involving laughter, criticism and teasing (Platt, 2008; Ruch, Altfreder, & Proyer, 2009). Considering that (a) verbal irony can be a humorous form of delivering a critical attitude (Garmendia, 2014), and (b) gelotophobes believe themselves to be a likely victim of humor and ridicule (which typically involves criticism), it can be hypothesized that gelotophobes have a disposition to mistake certain types of ironic verbal communication as literal, and also certain types of literal communication as ironic. Hence, there is reason to assume that the fear of being laughed at (gelotophobia) affects verbal irony comprehension, specifically when dealing with intended

compliments that are expressed via a mock criticism (false negative irony detection) and also when dealing with literal compliments (in terms of false positive irony detection).

There are several reasons to assume that gelotophobes may be prone to suspect that a sincere compliment is ironic, but also that an ironic criticism is literal, whereas non-gelotophobes do not. Firstly, gelotophobes are characterized by a strong sensitivity towards offense (Titze, 2009) and low self-esteem (see Ruch et al., 2014, for a review). Being convinced that they are deficient, they can be considered to have a tendency to expect being criticized by others, especially in a derisive way. Secondly, given that biting sarcasm can be employed to ridicule others, gelotophobes may suspect irony when receiving compliments because they have a bias to expect being ridiculed. Although the role of irony in bullying ridicule has not been investigated yet, irony may be used when putting down others in a derisive way. Hence, as a third explanation, gelotophobes may have been traumatized by events involving ridicule conveyed by ironic compliments and thereafter – with a paranoid tendency − view others as likely to address them with sarcasm or, in other words, may expect the worst when interacting with others. Fourth, considering that humor is often a by-product of irony (Garmendia, 2014), and that laughter is a typical response to humorous stimuli, gelotophobes can be assumed to generalize their paranoid sensitivity to anticipated ridicule (cf. Platt, Ruch, Hofmann, & Proyer, 2012). More specifically, it can be assumed that high scorers in gelotophobia believe that they are prone to encounter the very object of their fear also in the context of verbal irony: humor (which they believe to be put down by) and laughter (which they have a strong tendency to feel victimized by).

In order to gain a deeper understanding of why gelotophobes may be prone to take ironic criticisms as literal criticism, it may be instructive to take a closer look at ironic speech as a method of teasing, and gelotophobes' experiences of teasing situations. As Keltner, Capps, Ring, Young, and Heerey (2001) note, "[…] sarcasm and ironic utterances can be

forms of teasing when directed at another as a provocation or commentary" (p. 238). They view teasing to be central to human social life and support the notion that humor and play serves prosocial purposes such as socializing, flirting, entertaining and resolving conflicts. However, gelotophobes can be assumed to be oblivious to the good-natured character of teasing and mistake it as bullying ridicule. As Platt (2008) demonstrated, gelotophobes report more fear and shame than non-gelotophobes when dealing both with laughter situations in which they are addressed by a bullying-type of ridicule as well as in imagined situations involving good-natured teasing. Non-gelotophobic controls showed a distinct negative emotional reaction only to bullying ridicule. In other words, Platt's study showed that gelotophobes, when being addressed in a ridiculing or teasing manner, experience a stronger negative emotional reaction than non-gelotophobes. Furthermore, unlike non-gelotophobes, they do not distinguish between bullying ridicule and good-natured teasing. Understanding the mechanism and functions of teasing may help to explain why. Although they do not use the term *ridicule*, Boxer and Cortés-Conde (1997) postulate that teasing "[…] runs along a continuum of bonding to nipping to biting" (p. 279). They argue that teasing bonds in cases in which a humorous interaction (or *conversational joking* directed at someone present) is framed as *play,* while it *bites* when the playful character of the conversation is not recognized. According to Boxer and Cortés-Conde (1997), the target's construal of the teasing as *playful* moderates its emotional impact. Supporting this notion, Platt (2008) suggests that gelotophobes' undifferentiated negative response towards bullying ridicule and good-natured teasing originates from their failure to identify "the metamessage signal" conveyed by teasing that the "[…] interaction is playful, for fun and no harm is intended" (p. 122). However, Platt (2008) also notes that gelotophobes may recognize playful teasing signals but thereupon may not *trust* in the good-natured and fun character of the interaction. These considerations may help to further explain why gelotophobes can be assumed to take

ironic criticism as literal: They may not have developed a sense of being bonded to a social (peer) group but instead may have experienced traumatizing situations of being ostracized through ridicule in adolescence (cf. Titze, 2009). Therefore, they may distrust the cues signaling a playful and bonding interaction when being teased by benevolent ironic criticisms, taking them as literal criticisms.

Future development of the Test Of Verbal Irony Detection Aptitude (TOVIDA)

In order to make the TOVIDA usable in future studies investigating false alarms in irony detection, it would be necessary to construct a non-ironic criticism scale and a non-ironic praise scale. So far, the non-ironic items of the TOVIDA are used as fillers (or as a control scale in Study 2 of PART II). Future studies should employ the ironic criticism and ironic praise scale of the TOVIDA along with an extended list of non-ironic criticism and non-ironic praise items and develop two respective non-ironic scales. Furthermore, representative norms could be gathered to allow for normative comparisons of single individual's scores. These norms could help to identify high scorers and low scorers in irony detection aptitude without having to collect large samples for a certain investigation. Also, these norms could help to interpret test scores in an individual assessment setting. As to the psychometric properties, test-retest-correlations should verify the reliability of the scales and the stability of the constructs measured. Also, it must be seen as necessary to replicate the factorial structure in different (and preferably: large) samples, to further support the factorial validity indicated by PART II of the present thesis. Not least, the possibility to employ a more economic answer format should be explored, which can be seen as feasible when using the TOVIDA in the *irony-alert* mode of testing (i.e., when distracting test-takers from the intention of the test is not necessary).

Summarizing and organizing the constructs tested or assumed in the present thesis

One declared objective of the present thesis is to propose a theoretical model that summarizes and organizes the theoretical notions and empirical findings reported in the present thesis. The aim is to arrive at a psychological model of irony that accounts for individual differences in irony behavior. Following Eysenck's (1993) approach of deriving suggestions for a theory that accounts for the role of individual differences in creativity, the organization of the different constructs investigated in the present thesis was summarized in a theoretical model. Appendix A1 describes this relation.

As Appendix A1 summarizes, the argument is made that the success or the frequency of irony detection and use depend on four different groups of factors: (1) individual differences variables, i.e., ability variables (for example intelligence) and personality traits (such as psychoticism and the sense of humor). Furthermore, the model also comprises (2) mediating mechanisms that were assumed in the present thesis (e.g., *metamessage processing*; see section "Deriving a theory explaining successful irony reception" for details). To account for previous studies dealing with situational conditions that play a role in the interpretation of irony (such as the "common ground" or the quality of relationship between speaker and addressee, see Pexman & Zvaigzne, 2004, for an overview) or studies addressing cultural differences (such as between collectivist vs. individualist cultures, i.e., Rockwell & Theriot, 2001) that play a role in use, the model also includes (3) situational factors and (4) cultural factors. This may be seen as especially important if one takes into account that there may be personality-situation interactions: for example, *emotional stability* may come into play more prominently in ironic praise detection when self-involvement is high (vs. low, see general discussion in PART II of the present thesis).

This model describes that ironic communication is more (vs. less) frequent and more (vs. less) successful if the speaker and the listener score high (vs. low) in the abilities and

most of the traits listed, while some traits (i.e., seriousness, trait bad mood and gelotophobia) are supposed to be negatively associated to maximal and typical irony behavior. Situational and cultural factors may have both (a) main effects on the given irony behaviors, and (b) moderating effects on the association between ability and personality variables on the one hand and the irony behaviors on the other.

The variables listed are not claimed to be an exhaustive set, and some variables (such as humor creation ability and irony appreciation) that were not investigated in the present thesis have been integrated into the model. Furthermore, it is noteworthy that not all of the constructs were included based on sound empirical evidence found in the studies conducted for the present thesis; for some of them there were only correlational trends (i.e., agreeableness and culture, see PART II) which may indicate that substantial association may be found using different operationalizations (for example when using larger samples and more content valid instruments for their assessment). For other constructs listed in the model, preexisting findings indicate that they play a role in irony detection or use and should be included (e.g., creativity; Blouin & McKelvie, 2012; see also section "*Postformal thinking* and creativity in irony detection" in the General Introduction). Not least, the socially aversive personality traits delineated by the term "Dark Triad" were included based on the theoretical consideration that irony may be used as a means of aggression or disparagement of others, which may be a way subclinical psychopathy or narcissism may manifest in behavior (cf. Paulhus & Williams, 2002; see also PART III of the present thesis).

It needs to be noted that the model summarizing and organizing the constructs that the present thesis deals with must be seen as selective (as it is likely that relevant variables are not included). Furthermore, it needs to be seen whether (a) the findings integrated in this model are replicable and can be generalized, and (b) the assumptions made hold true when empirically tested. It is recommended to view the model as an initial summary and

organization of existing findings and relevant constructs which should be further elaborated, corrected, extended, and discussed in the light of future findings.

As a central aspect of this model, the *sense of irony* (see PART I) is postulated to explain interindividual variance in irony behavior. As described in the General Introduction, a theoretical construct consists of a system of *laws* that is referred to as a *nomological network* (Cronbach & Meehl, 1955). As Cronbach and Meehl (1995) point out, “"[l]earning more about" a theoretical construct is a matter of elaborating the nomological network in which it occurs, or of increasing the definiteness of the components.” (p. 290). To discuss the validity of this hypothetical construct, the “laws” by which the *sense of irony* manifests itself in irony behavior will be explored on a hypothetical basis in the next section.

Nomological network of the *sense of irony*

To discuss its validity, the *sense of irony* should be examined in terms of how, in its *nomological network,* (a) observable properties or quantities indicating the construct relate to each other, (b) the construct relates to observables; and (c) the construct relates to other theoretical constructs (cf. Cronbach & Meehl, 1955). The *sense of irony* is conceptualized as denoting (a) individuals’ aptitude to successfully deal with irony in terms of an ability component and (b) individuals’ tendency to enjoy, to like, or to seek irony in terms of a habitual preference component, both for the detection and the use of irony. Accordingly, individual differences variables from both the realm of personality and ability can be used for exploring the construct validity of the *sense of irony.*

How do observable properties or quantities relate to each other?

In PART I, the *sense of irony* was reasoned to include an *ability component* and a trait-like *habitual component*, which respectively refer to maximal irony behavior and typical irony behavior. Maximal irony behavior was defined as “a person’s capacity to produce or detect

irony" and it was reasoned that it "determines the upper limit of possible performance in the sense of ability" (Bruntsch et al., 2016, p. 31; i.e., PART I). The typical irony behavior, on the other hand, was defined as a person's habitual level of irony production and detection. This implies that individuals' typical levels of irony performance may deviate from their actual ability to perform irony. The observable properties of the *sense of irony* are summarized in Appendix A1. As Appendix A1 shows, maximal irony behavior comprises irony detection aptitude (as the reception ability) and irony creation ability (as the maximal production behavior). Typical irony behavior comprises irony use tendency (as the production preference) and irony appreciation (as the reception preference).

As a part of the nomological network of the *sense of irony*, a relationship between measures of irony *detection* aptitude and measures of irony use can be postulated. Importantly, this relationship can be expected to be moderated by irony preference. There may be individuals with a high aptitude to detect irony (as an indicator for the ability component of the *sense of irony*), who hardly ever use irony in daily life. On the level of *observable properties or quantities*, this consideration may lead to a possible operationalization for an empirical test of the assumption that there is an ability component and a habitual component in the sense of irony. Maximizing individuals' irony-alertness in an irony detection measure may serve as a way to measure the ability component. If participants were explicitly instructed to watch out for irony in an irony detection performance test (such as the TOVIDA), test scores would indicate how successful individuals deal with irony when instructed to do so. It can be reasoned that, if an irony use measure (such as the measures used in PART III) is employed without any mention of the relevance of irony, the ability component and the habitual component are confounded in the score. That is, the ability to successfully deal with irony is confounded with the habitual preference for irony. As an operationalization of irony preference, an irony *appreciation* measure can be employed as a moderator variable. If the

ability component and the habitual component of the *sense of irony* truly confound in a measure of irony use, one must expect that irony appreciation interacts with irony detection aptitude when explaining variance in irony use. This may be the case because irony appreciation (as an indicator for irony reception preference) cannot have an impact among individuals who do not have the ability to detect irony, but will manifest in irony use among individuals who are able to successfully deal with irony.

How does the* sense of irony *relate to observable quantities?

Since it includes an ability and a habitual component, the *sense of irony* must be viewed as a construct that manifests itself in both typical and maximal behavior. Accordingly, the sense of irony must be observable as an enduring tendency in a variety of different settings and contexts. In terms of measurement, the sense of irony is expected to be observable in the scores of performance measures, just as in measures of preference. It is noteworthy that, although the *sense of irony* is conceptualized as a unique construct, the observables indicating the *sense of irony* (e.g., irony detection performance) are also consequences of other traits and abilities. Appendix A1 summarizes relevant traits and abilities as predictors of irony behavior. As Appendix A1 shows, the *sense of irony* is listed as a construct that overlaps with both the ability and the personality domain of individual differences. Accordingly, the *sense of irony* is assumed to effect both maximal *and* typical irony behavior in terms of irony detection ability, irony appreciation, irony use tendency, and irony creation ability.

How do different theoretical constructs relate to the* sense of irony*?

Although the sense of irony must be seen as a narrower trait, it has a conceptual resemblance to constructs like creativity or the sense of humor (cf. Ruch, 1998). The *sense of irony is* conceptualized as a trait relating to personality and ability, in the same way that creativity is an individual characteristic associated with personality traits and ability (cf. Kandler et al., 2016). And just like creativity, the *sense of irony* is assumed to be an individual differences

variable that summarizes an ability component and a preference (or habitual) component. Furthermore, just like the sense of humor, the sense of irony manifests in different behavioral domains, such as the production and the reception of irony.

When defining the nomological network for the *sense of* irony, the relation between the construct and other theoretically relevant constructs is a crucial criterion for exploring the construct's validity. Usually, construct validity is studied in the course of evaluating the psychometric properties of measures for the assessment of a construct, especially when there is no suitable criterion measure of the construct (meaning one must resort to indirect measures). Importantly, construct validity is not a property of a test measure, because "[h]ere the trait or quality underlying the test is of central importance, rather than either the test behavior or the scores on the criteria" (Cronbach & Meehl, 1955, p. 282). Accordingly, the relation between the *sense of irony* and other theoretically relevant constructs can be seen as a way to define the nomological network of this construct. As in the *sense of irony,* an ability and a habitual component are assumed, the *sense of irony* is expected to relate to both ability and personality constructs. Since the observables indicating the *sense of irony* (e.g., irony detection performance) are also consequences of other traits and abilities, it almost seems natural that there is an overlap between measures for the *sense of irony* and measures of personality and ability variables relevant to irony use, such as the abilities and traits listed in Appendix A1. One important question that needs to be addressed is: How could one test whether the sense of irony is discrete from the constructs it converges with, such as the sense of humor and intelligence. The following section deals with this question.

Is the *sense of irony* a discrete construct?

As a central assumption, the *sense of irony* was conceptualized as a unique construct discrete from different constructs relevant to irony use (such as intelligence and humor-related traits,

see PART I of the present thesis). Therefore, future studies should test whether the *sense of irony* can predict irony behavior beyond known traits and ability. One suggested possible operationalization targeting maximal irony behavior could use a measure of irony detection aptitude as an indicator for the ability component of the *sense of irony,* and a measure of irony creation ability as the criterion. The irony creation performance measure should maximize irony alertness and employ an open-ended response format in order to assess individuals' capacity to generate irony, rather than the preference for ironic utterances when measuring irony use by employing pre-defined ironic utterances (see PART III). Using a hierarchical regression method, it should be possible to test whether the *sense of irony,* as measured, for example, with the Test of Verbal irony Detection Aptitude (TOVIDA, see PART II), can explain incremental variance in irony use beyond intelligence and humor creation ability (which can be measured, for example, with the Cartoon Punch line Production Test, cf. Köhler & Ruch, 1996).

One suggestion for testing whether the *sense of irony* explains incremental variance in *typical* irony behavior would be to use irony appreciation (as an indicator for the habitual component of the *sense of irony*) as a predictor*,* and a measure of irony use preference as the criterion. The irony use preference measure should minimize irony alertness so that individuals' habitual tendency to use irony can be assessed. Using a hierarchical regression method once more, it should be possible to test whether the *sense of irony,* as measured with a test assessing the appreciation of ironic stimuli, can explain incremental variance in irony use preference beyond broad personality dimensions and humor-related traits (i.e., foremost the ones found as predictors for irony use in PART III). If incremental variance is explained, this would be an indication for the *sense of irony* as a unique construct.

Deriving a theory explaining successful irony reception

By addressing an important aspect of irony behavior, the suggested theory of irony aims to explain "when and why" a listener successfully receives irony. However, this theory does not include situational or cultural factors, building instead on individual differences factors intervening in the communication process. So actually it aims to predict which listeners are more ready to detect irony successfully in terms of individual difference variables. As one of the crucial aspects of the suggested theory, it is proposed that irony has a *metamessage* (or also multiple metamessages). The metamessages of irony are assumed to be the "surplus" value of irony (beyond non-ironic communication) and typically go beyond the direct literal back-translation of an ironic utterance. The metamessages can be assumed as typically including a criticism toward the transgression of a social rule and running along a continuum from playful teasing to hostile ridicule.[17] As a prerequisite, the speaker must imply at least one meaningful metamessage by his ironic utterance in the sense of a purposeful communication contribution (see section "Grice's [1975] theory of conversational implicature and the cooperative principle" in the General Introduction of the present thesis). If this condition is satisfied, the communication process of ironic utterances can be assumed

[17] To illustrate: imagine that two friends watch a soccer match and friend A says "I bet you we will win this one! I studied the other's line-up, they have terrible players in the offense". When the other team scores the winning goal (meaning that the expectation fails), friend B might ironically comment to friend A "Terrible players in the offense!". The literal translation of this utterance on the "lowest" level could be "Excellent shot!". However, the metamessage of the ironic utterance goes beyond this translation, as friend B might be intending to teasingly ridicule his friend's false prediction by echoing his choice of words (i.e., by using the expression "terrible"). The possible critical metamessages might be, for example, "Don't boast with your analytical skills!", "Don't be too self-assured!", "Look at you, wannabe expert!", and so forth. *Multiple* metamessages (like in this example) with different contents may not only be admissible; they may actually be the very benefit that irony provides over literal communication. Also, different metamessages on different abstraction levels and in different facets of communication are thinkable. Accordingly, the metamessage may be encoded with factual information, an appeal, a self-revelation, and an assertion about a relationship (cf. Thun, 1981). For example, in the given situation a higher-level metamessage about the relationship between the speaker and the addressee may be: "I tease you in order to show you that I think we are close friends" (cf. Norrick, 1994).

to be mediated by individual differences variables facilitating the decoding of the metamessage(s).

As was established in PART II, more intelligent or emotionally stable individuals that are high in cheerfulness or low in bad mood, with high scores in benevolent or corrective humor, have a higher readiness to detect the irony in ironic praise (than less intelligent, emotionally unstable [i.e., neurotic], low-cheerful individuals, high in bad mood or low in benevolent or corrective humor). But *how* are these variables supposed to contribute to successful irony reception? As suggested in PART II, this may be the case because individuals with the respective scores are more able or ready to (a) reason and infer the metamessage of an ironic praise (i.e., *fluid intelligence*), (b) generate an easily interpreted mental "image" of the background of an ironic remark (i.e., *visual perception*), (c) reject the uttered criticism in what is literally said and recognize the more benevolent nature of what is ironically implied in the ironic praise (i.e., *emotional stability*), (d) take into account playful and humorous communicative intentions in terms of the processing of exhilarant stimuli (i.e., high *trait cheerfulness* and low *trait bad mood*), (e) have a smiling attitude towards the imperfections of life (e.g., human weakness) and know how to deal with them by using *benevolent humor*, and (f) expose transgressions of morally valued social rules by using irony with satirical metamessages in order to educate and better social others (i.e., the tendency to produce, to enjoy, and to make sense of *corrective humor*). Some of these assumptions about the mechanisms involved in irony detection will be taken when building a theory explaining successful irony reception in the next section.

The "Person-centered theory of irony reception"

The suggested theory, which shall be coined as the "Person-centered theory of irony reception" (accounting for its origin in the field of individual differences and for the sake of simplicity), posits that irony detection is facilitated by individual differences variables

(including the hypothesized *sense of* irony) that intervene in information processing when (1) rejecting the verbatim meaning of the ironic utterance (such as *emotional stability* and presumably also [low] *gelotophobia*) and (2) detecting that there is an implied *metamessage* behind the irony. The described assumed two-factorial process could be labeled as the "*reject-and-detect* axiom". In addition to this assumption, the individual differences variables suggested as possible mediators are assumed to help individuals to successfully process the humorously critical content of the metamessage of irony. Accordingly, the second important assumption of the suggested theory may be labeled as the "*metamessage processing* axiom", which posits that irony is only successfully received when the individual, due to his or her ability- and trait-dependent disposition, is ready to process its (humorously) critical *metamessage* effectively.

How does this theory fit in with the findings of PART II and PART III? As irony was postulated as humorous, it makes sense that the traits that were found as relevant for irony detection in PART II delineate tendencies toward jolly and compassionate humor behaviors (i.e., *cheerfulness* and *benevolent humor*). Furthermore, as irony is also theorized to be aggressive and critical (cf. Garmendia, 2014), it is coherent that, as a trait denoting the *combination* of humorous and critical behavior, *corrective humor* (as the tendency to satirically use "wit to ridicule vices, follies, abuses, and shortcomings with the intent of shaming individuals and groups into improvement", Ruch & Heintz, 2016a, p. 36) was found as a correlate of irony detection. As it is suggested in the section "The mixed-motive hypothesis of irony use", irony may be used interpersonally to be "*warm*" and "*cold*" at the same time (in terms of a *mixed-motive* communication). Accordingly, for the successful reception of irony, traits may be of special importance if they go along with the tendency to process the playful offense in the metamessage of irony (which is typically used when individuals ironically tease and ridicule others). In terms of the principle "It takes one to

know one", these might me the same traits that facilitate the use of irony as a double-edged vessel of humor and criticism, such as *katagelasticism* and the *aggressive humor style* (which were found among the correlates of irony use in PART III). This may be the case because the tendency to use these forms of humor (also in the form of irony) is likely to require the comprehension of these same forms of humor (i.e., as an aspect of the sense of humor overarching different behavioral domains).

To summarize, the "*Person-centered theory of irony reception*" suggested in this section makes assumptions as to (a) the *utilities* of irony (i.e., humor and aggression), (b) two processes jointly representing irony reception (i.e., rejecting the verbatim meaning and detecting the intended meaning), (c) the importance of the *metamessages* of irony, which typically criticize the transgression of social rules (but at the same time can include positive assertions about the relationship between the speaker and the addressee), and (d) the person-specific factors (i.e., traits and abilities) relating to the *utilities* of irony, relevant for one or both of the two processes jointly representing irony reception (i.e., "*reject* and *detect"*), and intervening in the processing of the metamessage. However, the axioms and many of the assumptions that the "Person-centered theory of irony reception" comprises have not been tested in the studies conducted for the present thesis. Hence, the theory still needs to be validated and should be refined by considering the findings of future works.

Application of the present findings

Humor training and dealing with intercultural challenges

As one possible application of the present findings, irony could be included in training interventions. Existing humor intervention programs do not target irony skills (see Ruch & McGhee, 2014, for an overview). Integrating irony as a humor technique could help to add new ways of being humorous to the behavioral repertoire coached in humor trainings.

Furthermore, it should help to identify individuals who are inclined to miscomprehend the intended meaning and could benefit from targeted interventions aiming to help develop irony skills. Adapting to cultural humor habits can be viewed as a necessity for employees of globally operating companies and institutions. Given that the assumption that the use of irony varies across cultures (and maybe also across sub-cultures) is valid (cf. Dress et al., 2008; Rockwell & Theriot, 2001), professional and social performance in an intercultural context may depend on an individual's opportunities to not only sensitize themselves to prevent missing irony, but also to prevent interpreting literal language as irony. As in the present thesis an instrument for the assessment of normally functioning adults' irony detection aptitude was developed and validated, this could help to design and tailor training interventions to the needs of the individual.

Using the TOVIDA in clinical studies

The findings of the current investigation suggest that future studies on irony detection aptitude should employ the two basic types of verbal irony that were investigated in this study, and to distinguish them using separate scores. This finding may not be restricted to the research question of who is inclined to mis-detect the intended meaning in ironic utterances among normally functioning adults: When working with clinical groups (i.e., patients with disorders or brain injuries that assumed as relevant to irony processing), deficits in irony detection aptitude may be more or less evident depending on which type of irony is being investigated. Taking into account that Langdon, Davies, and Coltheart (2002) found more pronounced results for banter (ironic praise) as compared to sarcasm (ironic criticism) when looking at schizophrenic patients' irony detection, it can be hypothesized that ironic praise (as the more challenging type of irony) may be more strongly affected by impaired or unusual functioning of cognitive and affective processing than ironic criticism (see PART II). Hence,

it may be worthwhile to also employ the TOVIDA for comparisons between patient groups and controls to test this assumption.

General Conclusion

The present thesis was first to target the question whether (a) systematic interindividual differences in irony performance can be found and (b) whether this interindividual variance can be explained by personality and ability variables. The results of the studies conducted for the present thesis paint a picture of *who* is able or inclined to detect or use irony. In dealing with these questions, it was possible to open up a new field of research for both personality and irony research by demonstrably linking the aptitude to detect irony and the tendency to use irony to ability and personality.

The present thesis introduced irony behavior as a phenomenon native to the realm of individual differences. As substantial relationship to general mental ability and personality traits were found, explaining irony behaviors should not remain a topic of experimental research (or the study of cultural groups or gender). The results of the studies conducted support the central claims of the present thesis in that they indicate that (a) there is systematic and measurable interindividual variance in irony detection and use (rather than merely negligible "noise"), and (b) this interindividual variance can be explained to a substantial degree by variables from the realm of ability and personality as two central domains of individual differences. While intelligence can be viewed as an important factor for maximal irony behavior (such as the ability to detect irony), typical irony behavior (such as the tendency to use irony) can be seen as linked to the affective, behavioral, cognitive, and motivational aspects summarized by personality traits.

As a secondary result of the studies in PART II and PART III, pre-existing assumptions about the special role of humor-related traits were put to the test, confirming that

(a) in irony detection, humor-related traits (i.e., cheerfulness or corrective humor) were predictive, while broad personality dimensions (i.e., the Big Five dimensions of personality) were only correlated in the case of emotional stability in the detection of ironic praise and (b) in irony use, humor-related traits explain incremental variance beyond broad personality dimensions (i.e., Eysenck's personality dimensions: psychoticism). This finding has an important implication for the *consequences* of the sense of humor. As temperamental traits constituting exhilaratability (e.g., trait cheerfulness) and also aspects of the sense of humor itself (e.g., benevolent and corrective humor) were found to be associated with individuals' aptitude to detect irony, the sense of humor can be seen as a trait relevant to successfully dealing with a range of different types of situations including irony in everyday life—such as in social and professional interactions (e.g., when flirting and teasing, or when debating and negotiating) or when consuming media products (for example: TV shows). The finding that humor-related traits correlate with irony detection aptitude highlights the importance of the sense of humor in everyday life (and, accordingly, the relevance of its investigation within psychological research). This is because having the disposition to miss out on the joke in humor (i.e., low scores in humor-related traits facilitating the experience of humor) can have significant consequences when this disposition is also associated with missing out on the meaning of irony encountered in daily life.

Ultimately, the findings of the studies conducted for the present thesis were integrated in a model that accounts for individual differences in irony behavior, and also includes the *sense of irony* as a newly introduced construct. The validity of the *sense of irony* was discussed by describing its nomological network and by listing possible designs for exploring its construct validity. The resulting model (see Appendix A1) was used as a basis for deriving a theory that explains successful irony reception. The suggested theory, which was coined as the "Person-centered theory of irony reception", aims to explain interindividual differences in

irony reception by making assumptions about (a) the *utilities* of irony (i.e., humor and aggression), (b) two processes jointly representing irony reception (i.e., rejecting the literal meaning and detecting the intended meaning), (c) the importance of the *metamessages* of irony, which typically criticize the transgression of social rules (but at the same time can include positive assertions about the relationship between the speaker and the addressee), and (d) the person-specific factors (i.e., traits and abilities) relating to the *utilities* of irony, relevant for one or both of the two processes jointly representing irony reception (i.e., the *reject-and-detect* aspect), as well as being beneficial for the processing of the metamessage. To conclude, the results of the theoretical and empirical work conducted for the present thesis demonstrate (a) associations between general mental ability and personality on the one hand, and irony detection on the other, and (b) associations between personality and irony use. Furthermore, the results validate newly introduced means of assessing irony detection aptitude and support pre-existing assumptions about the role of humor in irony behaviors. The findings have implications for different fields of irony research. For example, in the investigation of psychological and neuro-physiological correlates of irony, irony detection aptitude can be assessed in a more feasible and controlled fashion when the use of samples of normal functioning adults is made possible.

References

Ackerman, B. P. (1983). Form and function in children's understanding of ironic utterances. *Journal of Experimental Child Psychology*, *35*, 487–508. doi:10.1016/0022-0965(83)90023-1

Akimoto, Y., Sugiura, M., Yomogida, Y., Miyauchi, C. M., Miyazawa, S., & Kawashima, R. (2014). Irony comprehension: Social conceptual knowledge and emotional response. *Human Brain Mapping*, *35*, 1167–1178. doi:10.1002/hbm.22242

Anolli, L., Ciceri, R., & Infantino, M. G. (2002). From "blame by praise" to "praise by blame": Analysis of vocal patterns in ironic communication. *International Journal of Psychology*, *37*, 266–276. doi:10.1080/00207590244000106

Attardo, S. (2000). Irony as relevant inappropriateness. *Journal of Pragmatics*, *32*, 793–826. doi:10.1016/S0378-2166(99)00070-3

Attardo, S. (2001a). *Humorous Texts: A semantic and pragmatic analysis*. New York, US: Mouton De Gruyter.

Attardo, S. (2001b). Humor and irony in interaction: From mode adoption to failure of detection. In L. Anolli, R. Ciceri, & G. Riva (Eds.), *Say not to say: New perspectives on miscommunication* (pp. 159–179). Amsterdam, Netherlands: IOS Press.

Attardo, S., Eisterhold, J., Hay, J., & Poggi, I. (2003). Multimodal markers of irony and sarcasm. *Humor: International Journal of Humor Research*, *16*, 243–260. https://doi.org/10.1515/humr.2003.012

Averbeck, J. M., & Hample, D. (2008). Ironic message production: How and why we produce ironic messages. *Communication Monographs*, *75*, 396–410. doi:10.1080/03637750802512389

Bergson, H. L. (1924). *Laughter: An essay on the meaning of the comic*. New York, NY: MacMillan.

Blouin, P. S., & McKelvie, S. J. (2012). Postformal thinking as a predictor of creativity and of the identification and appreciation of irony and metaphor. *North American Journal of Psychology*, *14*, 39–50.

Bowes, A., & Katz, A. (2011). When Sarcasm Stings. *Discourse Processes*, *48*, 215–236. https://doi.org/10.1080/0163853X.2010.532757

Boxer, D., & Cortés-Conde, F. (1997). From bonding to biting: Conversational joking and identity display. *Journal of Pragmatics*, *27*, 275–294. doi:10.1016/S0378-2166(96)00031-8

Brüne, M., & Brüne-Cohrs, U. (2006). Theory of mind—evolution, ontogeny, brain mechanisms and psychopathology. *Neuroscience & Biobehavioral Reviews*, *30*, 437-455.

Bruntsch, R., Hofmann, J., & Ruch, W. (2016). Virgin soil in irony research: Personality, humor, and the "sense of irony." *Translational Issues in Psychological Science*, *2*, 25–34. doi:10.1037/tps0000054

Carroll, J. B. (1993). *Human cognitive abilities: A survey of factor-analytic studies*. Cambridge: Cambridge University Press.

Clark, H. H., & Gerrig, R. J. (1984). On the pretense theory of irony. *Journal of Experimental Psychology: General*, *113*, 121–126. doi:10.1037/0096-3445.113.1.121

Colston, H. L. (1997). Salting a wound or sugaring a pill: The pragmatic functions of ironic criticism. *Discourse Processes*, *23*, 25–45. doi:10.1080/01638539709544980

Colston, H. L., & Lee, S. Y. (2004). Gender differences in verbal irony use. *Metaphor and Symbol*, *19*, 289–306. doi:10.1207/s15327868ms1904_3

Colston, H. L., & Gibbs, R. W. J. (2007). A brief history of irony. In R. W. J. Gibbs & H. L. Colston (Eds.), *Irony in language and thought: A cognitive science reader.* (pp. 3–21). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Craik, K. H., Lampert, M. D., & Nelson, A. J. (1996). Sense of humor and styles of everyday humorous conduct. *Humor: International Journal of Humor Research*, *9*, 273–302.

Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281–302. doi:10.1037/h0040957

Dews, S., & Winner, E. (1995). Muting the meaning: A social function of irony. *Metaphor & Symbolic Activity*, *10*, 3–19. doi:10.1207/s15327868ms1001_2

Dews, S., Kaplan, J., & Winner, E. (1995). Why not say it directly? The social functions of irony. *Discourse Processes*, *19*, 347–367. doi:10.1080/01638539509544922

Dress, M. L., Kreuz, R. J., Link, K. E., & Caucci, G. M. (2008). Regional variation in the use of sarcasm. *Journal of Language and Social Psychology*, *27*, 71–85. doi:10.1177/0261927X07309512

Ducharme, L. J. (1994). Sarcasm and interactional politics. *Symbolic Interaction*, *17*, 51–62. doi:10.1525/si.1994.17.1.51

Dynel, M. (2014). Isn't it ironic? Defining the scope of humorous irony. *ResearchGate*, *27*. doi:10.1515/humor-2014-0096

Eysenck, H. J., & Eysenck, S. B. G. (1991). *Manual of the Eysenck Personality Scales (EPQ Adults)*. London, UK: Hodder and Stoughton.

Eysenck, H. J. (1992). The definition and measurement of psychoticism. *Personality and Individual Differences*, *13*, 757–785. doi:10.1016/0191-8869(92)90050-Y

Eysenck, H. J. (1993). Creativity and personality: Suggestions for a theory. *Psychological Inquiry*, *4*, 147–178. doi:10.1207/s15327965pli0403_1

Ferguson, M. A., & Ford, T. E. (2008). Disparagement humor: A theoretical and empirical review of psychoanalytic, superiority, and social identity theories. *Humor: International Journal of Humor Research*, *21*, 283–312. doi:10.1515/HUMOR.2008.014

Garmendia, J. (2014). The clash: Humor and critical attitude in verbal irony. *Humor: International Journal of Humor Research*, *27*, 641–659. doi:10.1515/humor-2014-0094

Gaudreau, G., Monetta, L., Macoir, J., Laforce, R. J., Poulin, S., & Hudon, C. (2013). Verbal irony comprehension in older adults with amnestic mild cognitive impairment. *Neuropsychology*, *27*, 702–712. doi:10.1037/a0034655

Gibbs, R. W. (1986). On the psycholinguistics of sarcasm. *Journal of Experimental Psychology: General*, *115*, 3–15. doi:10.1037/0096-3445.115.1.3

Gibbs, R. W. G., O'Brien, J. E., & Doolittle, S. (1995). Inferring meanings that are not intended: Speakers' intentions and irony comprehension. *Discourse Processes*, *20*, 187–203. doi:10.1080/01638539509544937

Gibbs, R. W. (2000). Irony in talk among friends. *Metaphor and Symbol*, *15*, 5–27. doi:10.1080/10926488.2000.9678862

Gibbs, R. W. J., & Colston, H. L. (2007). The future of irony studies. In R. W. J. Gibbs, H. L. Colston, R. W. J. (Ed) Gibbs, & H. L. (Ed) Colston (Eds.), *Irony in language and thought: A cognitive science reader.* (pp. 581–593). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Gibbs, R. J., Bryant, G. A., & Colston, H. L. (2014). Where is the humor in verbal irony? *Humor: International Journal of Humor Research*, *27*, 575–595. doi:10.1515/humor-2014-0106

Glenwright, M., & Pexman, P. M. (2010). Development of children's ability to distinguish sarcasm and verbal irony. *Journal of Child Language*, *37*, 429–451. doi:10.1017/S0305000909009520

Gómez, J.-C. (2008). The evolution of pretence: From intentional availability to intentional non-existence. *Mind & Language*, *23*, 586–606. doi:10.1111/j.1468-0017.2008.00353.x

Green, D. M., & Swets, J. A. (1966). *Signal detection theory and psychophysics*. Oxford, England: John Wiley.

Grice, H. P. (1975). Logic and conversation. In P. Cole & J. L. Morgan (Eds.), *Syntax and Semantics 3: Speech acts* (pp. 41–58). New York: Academic Press.

Groeben, N., Seemann, H., & Drinkmann, A. (1985). *Produktion und Rezeption von Ironie. Band II: Empirische Untersuchungen zu Bedingungen und Wirkungen ironischer Sprechakte* [Production and reception of irony. Volume 2: Empirical investigations of conditions and effects of ironic speech acts] (Tübinger Beiträge zur Linguistik; 279). Tübingen: Gunter Narr.

Groeben, N., & Scheele, B. (2003). Produktion von Ironie und Witz [Production of irony and wit]. In T. Herrmann, & J. Grabowski (Hrsg.[Eds.]), *Enzyklopädie der Psychologie: Sprachproduktion* [Encyclopedia of psychology: Speech production] (C, III, Bd. 1, S. 733–763). Göttingen, Germany: Hogrefe.

Hancock, J. T. (2004). Verbal irony use in face-to-face and computer-mediated conversations. *Journal of Language and Social Psychology*, *23*, 447–463. doi:10.1177/0261927X04269587

Happé, F. G. E. (1993). Communicative competence and theory of mind in autism: A test of relevance theory. *Cognition*, *48*, 101–119. doi:10.1016/0010-0277(93)90026-R

Haverkate, H. (1990). A speech act analysis of irony. *Journal of Pragmatics*, *14*, 77–109. doi:10.1016/0378-2166(90)90065-L

Hirsch, G. (2011). Between irony and humor: A pragmatic model. *Pragmatics & Cognition*, *19*, 530–561. doi:10.1075/pc.19.3.07hir

Holtgraves, T. (1997). Styles of language use: Individual and cultural variability in conversational indirectness. *Journal of Personality and Social Psychology*, *73*, 624–637. doi:10.1037/0022-3514.73.3.624

Ivanko, S. L., Pexman, P. M., & Olineck, K. M. (2004). How sarcastic are you? Individual differences and verbal irony. *Journal of Language and Social Psychology*, *23*, 244–271. doi:10.1177/0261927X04266809

Jorgensen, J. (1996). The functions of sarcastic irony in speech. *Journal of Pragmatics*, *26*, 613–634. doi:10.1016/0378-2166(95)00067-4

Kandler, C., Riemann, R., Angleitner, A., Spinath, F. M., Borkenau, P., & Penke, L. (2016). The nature of creativity: The roles of genetic factors, personality traits, cognitive abilities, and environmental sources. *Journal of Personality and Social Psychology*, *111*, 230–249. doi:10.1037/pspp0000087

Keltner, D., Capps, L., Kring, A. M., Young, R. C., & Heerey, E. A. (2001). Just teasing: A conceptual analysis and empirical review. *Psychological Bulletin*, *127*, 229–248. doi:10.1037/0033-2909.127.2.229

Köhler, G., & Ruch, W. (1996). Sources of variance in current sense of humor inventories: How much substance, how much method variance? *Humor: International Journal of Humor Research*, *9*, 363–397. doi:10.1515/humr.1996.9.3-4.363

Kotthoff, H. (2003). Responding to irony in different contexts: On cognition in conversation. *Journal of Pragmatics*, *35*, 1387–1411. doi:10.1016/S0378-2166(02)00182-0

Kreuz, R. J., & Glucksberg, S. (1989). How to be sarcastic: The echoic reminder theory of verbal irony. *Journal of Experimental Psychology: General*, *118*, 374–386. doi:10.1037/0096-3445.118.4.374

Kreuz, R. J., Long, D. L., & Church, M. B. (1991). On being ironic: Pragmatic and mnemonic implications. *Metaphor and Symbolic Activity*, *6*, 149–162. doi:10.1207/s15327868ms0603_1

Kreuz, R. J., & Roberts, R. M. (1995). Two cues for verbal irony: Hyperbole and the ironic tone of voice. *Metaphor and Symbolic Activity*, *10*, 21–31. doi:10.1207/s15327868ms1001_3

Kreuz, R. J. (2000). The production and processing of verbal irony. *Metaphor and Symbol*, *15*, 99–107. doi:10.1080/10926488.2000.9678867

Kreuz, R. J., & Link, K. E. (2002). Asymmetries in the use of verbal irony. *Journal of Language and Social Psychology*, *21*, 127–143. doi:10.1177/02627X02021002002

Kreuzpointner, L., Lukesch, H. & Horn, W. (2013). *LPS-2. Leistungsprüfsystem 2* [Achievement Measurement System 2]. Göttingen: Hogrefe.

Kumon-Nakamura, S., Glucksberg, S., & Brown, M. (1995). How about another piece of pie: The allusional pretense theory of discourse irony. *Journal of Experimental Psychology: General*, *124*, 3–21. doi:10.1037/0096-3445.124.1.3

Langdon, R., Davies, M., & Coltheart, M. (2002). Understanding minds and understanding communicated meanings in schizophrenia. *Mind & Language*, *17*, 68–104. doi:10.1111/1468-0017.00190

Langdon, R., & Coltheart, M. (2004). Recognition of metaphor and irony in young adults: the impact of schizotypal personality traits. *Psychiatry Research*, *125*, 9–20. doi:10.1016/j.psychres.2003.10.005

Leggitt, J. S., & Gibbs, R. W. (2000). Emotional reactions to verbal irony. *Discourse Processes*, *29*, 1–24. doi:10.1207/S15326950dp2901_1

Legree, P. J. (1995). Evidence for an oblique social intelligence factor established with a Likert-based testing procedure. *Intelligence*, *21*, 247–266. doi:10.1016/0160-2896(95)90016-0

Littman, D. C., & Mey, J. L. (1991). The nature of irony: Toward a computational model of irony. *Journal of Pragmatics*, *15*, 131–151. doi:10.1016/0378-2166(91)90057-5

Lucariello, J. (1994). Situational irony: A concept of events gone awry. *Journal of Experimental Psychology: General*, *123*, 129–145. doi:10.1037/0096-3445.123.2.129

Martin, R. A. (1998). Approaches to the sense of humor: A historical review. In W. Ruch (Ed.), *The sense of humor: Explorations of a personality characteristic.* (pp. 15–60). Berlin, Germany: Walter de Gruyter & Co.

Martin, R. A., Puhlik-Doris, P., Larsen, G., Gray, J., & Weir, K. (2003). Individual differences in uses of humor and their relation to psychological well-being: Development of the Humor Styles Questionnaire. *Journal of Research in Personality*, *37*, 48–75. doi:10.1016/S0092-6566(02)00534-2

Martin, I., & McDonald, S. (2004). An exploration of causes of non-literal language problems in individuals with Asperger syndrome. *Journal of Autism and Developmental Disorders*, *34*, 311–328. doi:10.1023/B:JADD.0000029553.52889.15

Matthews, J. K., Hancock, J. T., & Dunham, P. J. (2006). The roles of politeness and humor in the asymmetry of affect in verbal irony. *Discourse Processes*, *41*, 3–24. doi:10.1207/s15326950dp4101_2

McDonald, S., & Pearce, S. (1996). Clinical insights into Pragmatic Theory: frontal lobe deficits and sarcasm. *Brain and Language*, *53*, 81–104. doi:10.1006/brln.1996.0038

McDonald, S., Bornhofen, C., Shum, D., Long, E., Saunders, C., & Neulinger, K. (2006). Reliability and validity of The Awareness of Social Inference Test (TASIT): A clinical test of social perception. *Disability and Rehabilitation*, *28*, 1529–1542. doi:10.1080/09638280600646185

Milanowicz, A. (2013). Irony as a means of perception through communication channels. emotions, attitude and IQ related to irony across gender. *Psychology of Language and Communication*, *17*(2), 115–132.

Mitchley, N. J., Barber, J., Gray, J. M., Brooks, D. N., & Livingston, M. G. (1998). Comprehension of irony in Schizophrenia. *Cognitive Neuropsychiatry*, *3*, 127–138. doi:10.1080/135468098396206

Monetta, L., Grindrod, C. M., & Pell, M. D. (2009). Irony comprehension and theory of mind deficits in patients with Parkinson's disease. *Cortex*, *45*, 972–981. doi:10.1016/j.cortex.2009.02.021

Norrick, N. R. (1994). Involvement and joking in conversation. *Journal of Pragmatics*, *22*, 409–430. doi:10.1016/0378-2166(94)90117-1

Oomen, U. (1983). Ironische Äusserungen: Syntax–Semantik–Pragmatik. *Zeitschrift für Germanistische Linguistik, 11*, 22-38. doi:10.1515/zfgl.1983.11.1.22

Ostendorf, F. (1990). *Sprache und Persönlichkeitsstruktur. Zur Validität des Fünf-Faktoren-Modells der Persönlichkeit* [Language and personality. On the validity of the five-factor model of personality]. Regensburg, Germany: Roderer.

Paulhus, D. L., & Williams, K. M. (2002). The Dark Triad of personality: Narcissism, Machiavellianism, and psychopathy. *Journal of Research in Personality*, *36*, 556–563. doi:10.1016/S0092-6566(02)00505-6

Pexman, P. M., & Olineck, K. M. (2002). Does sarcasm always sting? Investigating the impact of ironic insults and ironic compliments. *Discourse Processes*, *33*, 199–217. doi:doi.org/10.1207/S15326950DP3303_1

Pexman, P. M., & Zvaigzne, M. T. (2004). Does irony go better with friends? *Metaphor and Symbol*, *19*, 143–163. doi:10.1207/s15327868ms1902_3

Pexman, P. M. (2005). Social Factors in the Interpretation of Verbal Irony: The Roles of Speaker and Listener Characteristics. In H. L. Colston & A. N. Katz (Eds.), *Figurative language comprehension: Social and cultural influences* (pp. 209–232). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Platt, T. (2008). Emotional responses to ridicule and teasing: Should gelotophobes react differently? *Humor: International Journal of Humor Research*, *21*, 105–128. doi:10.1515/HUMOR.2008.005

Platt, T., Ruch, W., Hofmann, J., & Proyer, R. T. (2012). Extreme fear of being laughed at: Components of gelotophobia. *The Israeli Journal of Humor Research: An International Journal, 1*, 86-106.

Rapp, A. M., & Wild, B. (2011). Nonliteral language in Alzheimer dementia: A review. *Journal of the International Neuropsychological Society*, *17*, 207–218. doi:10.1017/S1355617710001682

Raskin, V. (1985). *Semantic mechanisms of humor*. Dordrecht, Netherlands: D. Reidel.

Renner, K.-H., Enz, S., Friedel, H., Merzbacher, G., & Laux, L. (2008). Doing as if: The histrionic self-presentation style. *Journal of Research in Personality*, *42*, 1303–1322. doi:10.1016/j.jrp.2008.04.005

Renner, K.-H., & Heydasch, T. (2010). Performing humor: On the relations between self-presentation styles, gelotophobia, gelotophilia, and katagelasticism. *Psychological Test and Assessment Modeling*, *52*, 171–190.

Roberts, R. M., & Kreuz, R. J. (1994). Why do people use figurative language? *Psychological Science*, *5*, 159–163. doi:10.1111/j.1467-9280.1994.tb00653.x

Rockwell, P., & Theriot, E. M. (2001). Culture, gender, and gender mix in encoders of sarcasm: A self‐assessment analysis. *Communication Research Reports*, *18*, 44–52. doi:10.1080/08824090109384781

Ruch, W., & Deckers, L. (1993). Do extraverts 'like to laugh'?: An analysis of the Situational Humor Response Questionnaire (SHRQ). *European Journal of Personality*, *7*, 211–220. doi:10.1002/per.2410070402

Ruch, W., Köhler, G., & Van Thriel, C. (1996). Assessing the "humorous temperament": Construction of the facet and standard trait forms of the State-Trait-Cheerfulness-Inventory–STCI. *Humor: International Journal of Humor Research*, *9*, 303–339. doi:10.1515/humr.1996.9.3-4.303

Ruch, W. (1998). Foreword and overview. The sense of humor: A new look at an old concept. In W. Ruch (Ed.), *The sense of humor: Explorations of a personality characteristic* (pp. 3–14). Berlin, Germany: Mouton de Gruyter.

Ruch, W. (1999). Die revidierte Fassung des Eysenck Personality Questionnaire und die Konstruktion des deutschen EPQ-R bzw. EPQ-RK. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *20*, 1–24. doi:10.1024//0170-1789.20.1.1

Ruch, W. (2004). Humor. In C. P. Peterson, & M. E. P. Seligman (Eds.), *Character strengths and virtues: A handbook and classification* (pp. 583-598). Washington, DC: American Psychological Association.

Ruch, W. (2008). The psychology of humor. In V. Raskin (Ed.), *A primer of humor* (pp. 17-100). Berlin, Germany: Mouton de Gruyter.

Ruch, W., & Proyer, R. T. (2008). Who is gelotophobic? Assessment criteria for the fear of being laughed at. *Swiss Journal of Psychology, 67*, 19–27. doi:10.1024/1421-0185.67.1.19

Ruch, W., Altfreder, O., & Proyer, R. T. (2009). How do gelotophobes interpret laughter in ambiguous situations? An experimental validation of the concept. *Humor: International Journal of Humor Research*, *22*, 63–89. doi:10.1515/HUMR.2009.004

Ruch, W., & Proyer, R. T. (2009). Extending the study of gelotophobia: On gelotophiles and katagelasticists. *Humor: International Journal of Humor Research*, *22*, 183–212. doi:10.1515/HUMR.2009.009

Ruch, W., & Hofmann, J. (2012). A temperament approach to humor. In P. Gremigni (Ed.), *Humor and health promotion* (pp. 79–113). Hauppauge, New York, NY: Nova Science Publishers.

Ruch, W. F., & Heintz, S. (2014). Humour styles, personality and psychological well-being: What's humour got to do with it? *The European Journal of Humour Research*, *1*, 1–24. doi:10.7592/EJHR2013.1.4.ruch

Ruch, W., Hofmann, J., Platt, T., & Proyer, R. (2014). The state-of-the art in gelotophobia research: A review and some theoretical extensions. *Humor: International Journal of Humor Research*, *27*, 23–45. doi:10.1515/humor-2013-0046

Ruch, W., & McGhee, P. E. (2014). Humor intervention programs. In A. C. Parks & S. M. Schueller (Eds.), *Handbook of Positive Psychological Interventions* (pp. 179-193). Oxford: Wiley-Blackwell.

Ruch, W., & Heintz, S. (2016a). The virtue gap in humor: Exploring benevolent and corrective humor. *Translational Issues in Psychological Science*, *2*, 35–45. doi:10.1037/tps0000063

Ruch, W., & Heintz, S. (2016b). The German version of the Humor Styles Questionnaire: Psychometric properties and overlap with other styles of humor. *Europe's Journal of Psychology*, *12*, 434–455. doi10.5964/ejop.v12i3.1116

Ruch, W., & Heintz, S. (2017). Experimentally manipulating items informs on the (limited) construct and criterion validity of the Humor Styles Questionnaire. *Frontiers in Psychology*, *8*, 616. doi:10.3389/fpsyg.2017.00616

Rydell, S. T., & Rosen, E. (1966). Measurement and some correlates of need-cognition. *Psychological Reports*, *19*, 139–165. doi:10.2466/pr0.1966.19.1.139

Schaafsma, S. M., Pfaff, D. W., Spunt, R. P., & Adolphs, R. (2015). Deconstructing and reconstructing theory of mind. *Trends in Cognitive Sciences*, *19*, 65–72. doi:10.1016/j.tics.2014.11.007

Schallberger, U. & Venetz, M. (1999). Kurzversionen des MRS-Inventars von Ostendorf (1990) zur Erfassung der fünf "grossen" Persönlichkeitsfaktoren. [Short version of the MRS inventory by Ostendorf (1990) for the assessement of the “big” five personality factors]. Berichte aus der Abteilung Angewandte Psychologie, Nr. 30. Zürich: Psychologisches Institut der Universität Zürich [Reports from the section of applied psychology, No. 30. Zurich: Department for Psychology, University of Zurich].

Schmidt-Hidding, W. (1963). *Europäische Schlüsselwörter. Band I: Humor und Witz.* [European key terms. Volume I: Humor and wit]. Munich, Germany: Huber.

Schwoebel, J., Dews, S., Winner, E., & Srinivas, K. (2000). Obligatory processing of the literal meaning of ironic utterances: Further evidence. *Metaphor and Symbol*, *15*, 47–61. doi:10.1080/10926488.2000.9678864

Shott, S. (1979). Emotion and social life: A symbolic interactionist analysis. *American Journal of Sociology*, *84*, 1317–1334.

Sperber, D., & Wilson, D. (1981). Irony and the use-mention distinction. In P. Cole (Ed.), *Radical pragmatics* (pp. 295-318). New York: Academic Press.

Suls, J. M. (1972). A two-stage model for the appreciation of jokes and cartoons: An information-processing analysis. In J. H. Goldstein & P. E. McGhee (Eds.), *The psychology of humor: Theoretical perspectives and empirical issues* (pp. 81–100), New York, NY: Academic Press.

Tangney, J. P., Miller, R. S., Flicker, L., & Barlow, D. H. (1996). Are shame, guilt, and embarrassment distinct emotions? *Journal of Personality and Social Psychology*, *70*, 1256–1269. doi:10.1037/0022-3514.70.6.1256

Thun, F. S. von. (1981). *Miteinander reden, Vol. 1: Störungen und Klärungen - Allgemeine Psychologie der Kommunikation* [Talk to each other, Vol 1: Interference and clarification – General Psychology of communication]. Reinbek bei Hamburg: Rowohlt.

Titze, M. (2009). Gelotophobia: The fear of being laughed at. *Humor: International Journal of Humor Research*, *22*, 27–48. doi:10.1515/HUMR.2009.002

Utsumi, A. (2000). Verbal irony as implicit display of ironic environment: Distinguishing ironic utterances from nonirony. *Journal of Pragmatics*, *32*, 1777–1806. doi:10.1016/S0378-2166(99)00116-2

Varga, E., Schnell, Z., Tényi, T., Németh, N., Simon, M., Hajnal, A., … Herold, R. (2014). Compensatory effect of general cognitive skills on non-literal language processing in schizophrenia: A preliminary study. *Journal of Neurolinguistics*, *29*, 1–16. doi:10.1016/j.jneuroling.2014.01.001

Veselka, L., Schermer, J. A., Martin, R. A., & Vernon, P. A. (2010). Relations between humor styles and the Dark Triad traits of personality. *Personality and Individual Differences*, *48*, 772–774. doi:10.1016/j.paid.2010.01.017

Whiten, A. (1992). Mind reading, pretence and imitation in monkeys and apes. *Behavioral and Brain Sciences*, *15*, 170–171. doi:10.1017/S0140525X00068187

Wilt, J., & Revelle, W. (2015). Affect, behaviour, cognition and desire in the Big Five: An analysis of item content and structure. *European Journal of Personality*, *29*, 478–497. doi:10.1002/per.2002

Winner, E., Brownell, H., Happé, F., Blum, A., & Pincus, D. (1998). Distinguishing lies from jokes: theory of mind deficits and discourse interpretation in right hemisphere brain-damaged patients. *Brain and Language*, *62*, 89–106. doi:10.1006/brln.1997.1889

Zillmann, D. (1983). Disparagement humor. In P. E. McGhee & J. H. Goldstein (Eds.), *Handbook of humor research: Vol. 1. Basic Issues* (pp. 85–107). New York, NY: Springer.

Ziv, I., Leiser, D., & Levine, J. (2011). Social cognition in schizophrenia: Cognitive and affective factors. *Cognitive Neuropsychiatry*, *16*, 71–91. doi:10.1080/13546805.2010.492693

Ehrenwort

Hiermit erkläre ich, dass die Dissertation von mir selbst ohne unerlaubte Beihilfe verfasst worden ist.

Zürich, am

_____

Ort und Datum Unterschrift

Curriculum Vitae

Name: Richard Bruntsch

Date of birth: 17.05.1983

Education

University and academic degree

2004-2012	Psychology, University of Heidelberg (Ruprecht-Karls-Universität Heidelberg, degree: "Diplom-Psychologe", predecessor of/equivalent to MSc, final grade: 1.4 = "very good")
2007-2008	Semester at the School of Psychology, University of Southampton, England, with ERASMUS (European Community Action Scheme for the Mobility of University Students) program
2006	Passing qualification for advanced studies in psychology, Ruprecht-Karls-Universität Heidelberg ("Vordiplom", grade: 1.5 = "very good")
2003-2004	Sociology, Ruprecht-Karls-Universität Heidelberg

Secondary

1993-2002	Ludwig-Georgs-Gymnasium Darmstadt, Abitur (final grade: 1.6 = "good")

Professional Training

2015-	PhD student at the University of Zurich
2013-	Assistant at the Section of Personality and Assessment, Department of Psychology, University of Zürich

University Teaching and Thesis Supervision

2017	Lecturing "Clinical Assessment" (1 session) in MSc lecture "Psychological Assessment"
2014-	Teaching MSc students in the seminars: "Einführung in die Testkonstruktion: Theorie und Praxis" [Introduction to test construction: theory and practice] at the Department of Psychology, University of Zürich "Praktische Interventionen 2: Anwendungsgebiete und praktische Übungen zu psychologisch-diagnostischen Verfahren" [Fields of application of and practical training on

psychological diagnostic methods] at the Department of Psychology, University of Zürich

Teaching BSc-students in the seminar: "Experimentalpsychologisches Praktikum" [Experimental practice seminar] at the Department of Psychology, University of Zürich

2014- Teaching assistant at the Swiss distance learning university in personality psychology (www.fernuni.ch)

2015- Supervision of MSc-theses at the Department of Psychology, University of Zürich:
"The association of gelotophobia and irony detection" (Isabelle Brunner, BSc), "Evaluating a novel performance test for the assessment of irony detection" (Jasmine Fong, BSc), "The role of humor and personality in irony use" (Kristina Rettich, BSc), "Is there a sense of irony?" (Fabian Langensteiner, BSc)

2014- Supervision of BSc theses at the Department of Psychology, University of Zürich:
"Individual differences in the comprehension and use of irony and sarcasm" (Stefan Schönholzer),
"Is resiliency a matter of personality?" (Christine Wolfer),
"Self-esteem: aggregation of domain-specific self-concepts?" (Vincy Chennamparampil),
"How are personality and dysfunction related" (Barbara Boller),
"Personality and intoxicant abuse" (Cecilia Longo),
"Theory of Mind: An individual differences variable?" (Jaqueline Ritzmann),
"Shyness in the context of personality" (Jan Juchli).

Research

Research interests: Study of interindividual differences in irony detection and use, personality and character in the moral domain, development of psychometric measures (e.g., for the assessment of irony detection aptitude, gelotophobia, and cynicism)

Academic awards

2015 "Best Poster" award at the 13th European Conference on Psychological Assessment (ECPA13), Zurich, Switzerland, for the contribution: Bruntsch, R., & Ruch, W. (2015, July). Evaluating the rationale of a novel irony performance test.

Professional Memberships

2015- Member of the European Association of Psychological Assessment (EAPA, http://www.eapa-homepage.org)

2014-	Member of the International Society for Humor Studies (ISHS, www.humorstudies.org)
2014-	Founding member of the Swiss Society for Positive Psychology (SWIPPA, www.swippa.ch)
	Administrative academic functions
2014-2015	Member of the organizing committee for the 13th European Conference on Psychological Assessment (ECPA13) of the European Association for Psychological Assessment (EAPA), 22 to 25 July 2015, Zürich, Switzerland
2015	Invited peer reviewer for Translational Issues in Psychological Science (TPS)
	Under-graduate internships
2011	Research internship at Zentralinstitut für Seelische Gesundheit (ZI) Mannheim, AG Imaging in der Psychiatrie
2009	Internship at Zentralinstitut für Seelische Gesundheit (ZI) Mannheim, Klinik für Psychosomatik und Psychotherapeutische Medizin
	Undergraduate Research and Dissemination Experience
2009-2010	Student helper in the section of Social Psychology of the University of Heidelberg
2008	Student helper in the research project “SOPHOS” on social phobia at Medizinische Fakultät Heidelberg, Klinik für psychosomatische und allgemeine klinische Medizin
2007-2008	Student helper in in the research projects “SENSO-AGE” and “INSEL” of the section of Psychological Research on Ageing of the University of Heidelberg
	Community service
2002-2003	Care of handicapped persons at „Club Behinderter und ihrer Freunde in Darmstadt und Umgebung e.V.“

Publications

Journal articles

Bruntsch, R., & Ruch, W. (2017). Studying irony detection beyond ironic criticism: Let's include ironic praise. *Frontiers in Psychology*, *8*:606. doi:10.3389/fpsyg.2017.00606

Bruntsch, R., & Ruch, W. (2017). The role of humor-related traits and broad personality dimensions in irony use. *Personality and Individual Differences*, *112*, 139–143. doi:10.1016/j.paid.2017.03.004

Ruch, W., **Bruntsch, R.**, & Wagner, L. (2017). The role of character traits in economic games. *Personality and Individual Differences*, *108*, 186–190. doi:10.1016/j.paid.2016.12.007

Bruntsch, R., Hofmann, J., & Ruch, W. (2016). Virgin soil in irony research: Personality, humor, and the "sense of irony." *Translational Issues in Psychological Science*, *2*, 25–34. doi:10.1037/tps0000054

Posters presented at professional meetings

Ruch, W., **Bruntsch, R.**, & Wagner, L. (2016, May). Character strengths predict decisions in economic games over and above personality traits. Poster presented at the 13th MaDoKo of the Department of Psychology at the University of Zurich, Zurich, Switzerland.

Bruntsch, R., & Ruch, W. (2015, July). Evaluating the rationale of a novel irony performance test. Poster presented at the 13th European Conference on Psychological Assessment (ECPA13), Zurich, Switzerland.

Bruntsch, R., Platt, T., & Ruch, W. (2014, May). Psychometric properties, structure, and concurrent validity of a picture-based test for the assessment of gelotophobia. Poster presented at the 11th LiMaDoKo of the Department of Psychology at the University of Zurich, Zurich, Switzerland.

Talks delivered at professional meetings

Bruntsch, R., & Ruch, W. (2016). Studying correlates of irony detection performance: Do not hide the measurement intention!. Talk at the 28th International Society for Humor Studies (ISHS) Conference, June 27th-July 1st 2016, Dublin, Ireland, UK.

Bruntsch, R., Ruch, W., & Platt, T. (2014). Evaluation of a picture-based test for the assessment of gelotophobia (Picture-Geloph). Talk at the 26th International Society for Humor Studies (ISHS) Conference, July 7th-11th 2014, Utrecht, The Netherlands.

Bruntsch, R., Ruch, W., & Platt, T. (2014). Proposing a short form of the Picture-Geloph. Talk at the 14th International Summer School and Symposium on Humour and Laughter: Theory, Research and Applications, July 14th-19th 2014, Sheffield, United Kingdom

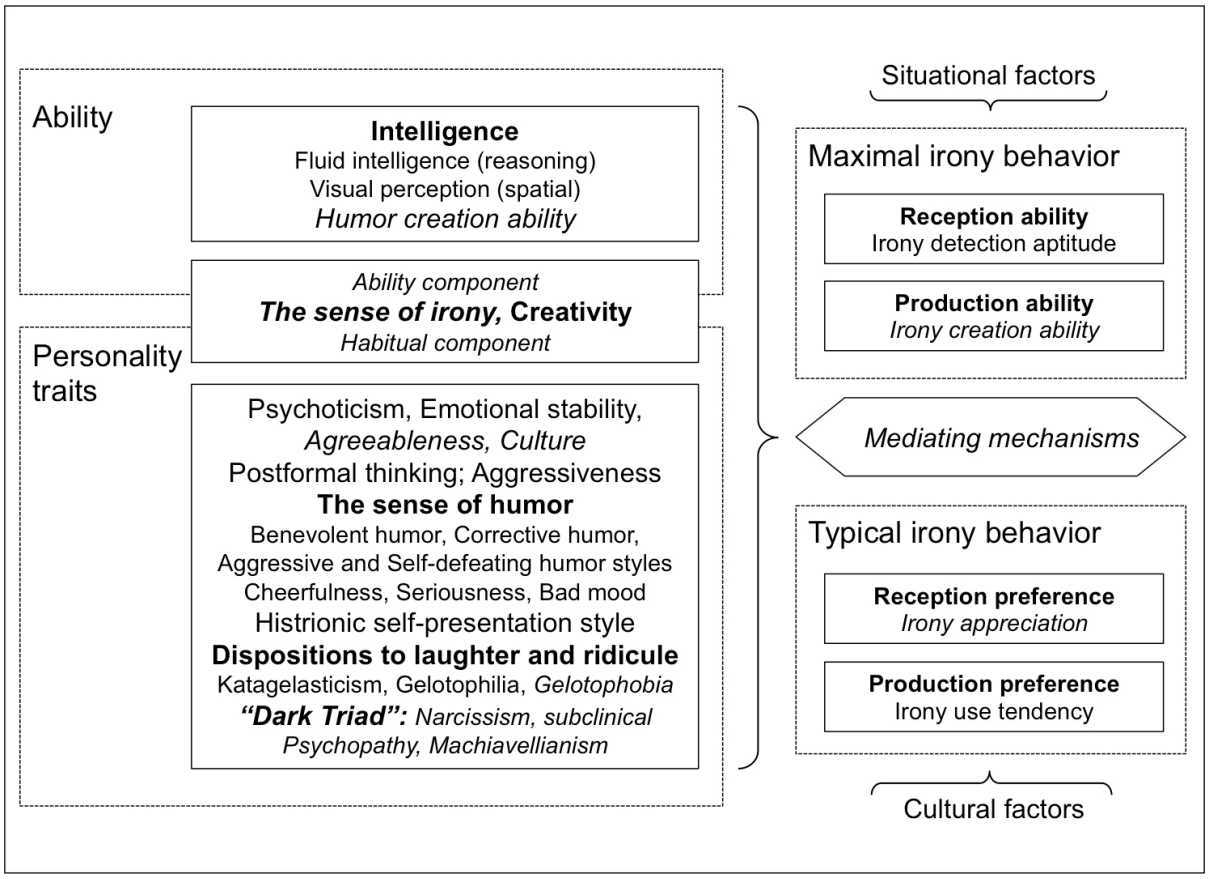
Invited talks/workshops

Bruntsch, R. (2016). "Ironie und Humor in der Pädagogik" [Irony and humor in education]. Talk for teachers of the Kantonschule Zürcher Unterland, Schloss Marbach, Germany, November 18th 2016.

Bruntsch, R. (2015). "Verbale Ironie besser verstehen und übersetzen lernen" [Learning more about verbal irony and its translation]. Workshop for Berufsvereinigung GebärdensprachdolmetscherInnen Deutschschweiz, Zürich, Switzerland, November 21st 2015.

Appendix

Appendix A1: Suggested model summarizing and organizing constructs and factors featured in the present thesis



Variables that are assumed as relevant or found as relevant in the studies conducted for the present thesis or previous studies. Variables for which no substantial empirical support has been found yet are printed in italics. For the sake of simplicity, interactions between ability and personality variables as well as situational factors and the mediators assumed by the "Person-centered theory of irony reception" (e.g., metamessage decoding) are not explicated.

Appendix A2: Ironic praise sample item

The following sample item of the TOVIDA given below is translated from German language. The statements for the appraisal of the situation were rated on a four-point scale as to how much the sentences apply to the situation (1= "does not apply at all", 2="rather does not apply", 3= "rather applies", 4= "fully applies"). Statements printed in bold were used as (inversed) indicators of irony detection in the studies and averaged to build the item score. Three of the appraisal statements were designed as distractors for each item.

Ironic praise sample item (IP6):

Situation:

Christian has invited three friends over for dinner. He prepares a meal trying a new recipe. Sitting at the table starting to eat, Julia asks for the saltshaker. Christian immediately apologizes that he could not salt the food to taste as he has a cold and cannot taste properly. This is when Julia says: "You are right, the food is inedible!".

Instruction: *Please indicate how much you think that each of the following statements applies to the situation.*

Statements:

1. The saltshaker is not within Julia's reach.

2. Christian behaves like a typical male.

3. Julia is having a good time.

***4. Christian will feel bad because of Julia's final utterance.** (-)*

***5. Julia will not finish her plate without more salt.** (-)*

6. Christian is making up excuses